

PLACE IN RETURN BOX to remove this checkout from your record.
TO AVOID FINES return on or before date due.
MAY BE RECALLED with earlier due date if requested.

DATE DUE	DATE DUE	DATE DUE

**BIOMOLECULAR ENGINEERING OF siRNA
THERAPEUTICS**

By

JOSEPH A. GREDELL

A DISSERTATION

**Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of**

DOCTOR OF PHILOSOPHY

Chemical Engineering

2009

ABSTRACT

BIOMOLECULAR ENGINEERING OF siRNA THERAPEUTICS

By

JOSEPH A. GREDELL

RNA interference (RNAi) provides a powerful means for regulating gene expression. Although it is a relatively recent discovery, it has already proven useful in identification of gene/protein functions and is in the process of being utilized for disease treatment in humans. Continued use of this pathway for a variety of applications would benefit from a more thorough understanding of the role of the initiator molecules of RNAi, small interfering RNA (siRNA). As the efficacy of siRNAs that mediate RNAi is known to vary greatly, the aim of this thesis was to study three steps of the pathway that pose barriers to identifying and developing highly active siRNA molecules for use as therapeutic agents.

The first of these factors investigated in this research was the effect of secondary structure within the mRNA site targeted by siRNA. The results obtained from both experimental and computational studies show that sites with extensive mRNA secondary structure were less susceptible to silencing, while those containing unpaired nucleotides at either the 5'- or 3'-end were generally more amenable. Similar observations were made when taking into account siRNA guide strand structure. Taken together, there is a correlation between RNA structure and silencing efficiency that ultimately can be included in existing and future siRNA selection algorithms for the improved identification of active siRNAs, thus reducing the number of sequences to be tested before selecting one capable of significantly silencing the target gene.

The second aspect of RNAi explored here was the influence of siRNA sequence and hybridization stability on recognition by the TAR RNA binding protein (TRBP). This protein is part of the RNA induced silencing complex (RISC) responsible for targeting and cleaving mRNAs to achieve silencing. It was found that TRBP can detect the overall and relative stability of the two ends of the siRNA by interacting primarily with the more stable end, and that the interaction is similarly reflected when binding to single-stranded RNAs alone. Additionally, TRBP interactions can be altered through the inclusion of mismatches within the siRNA sequence or DNA substitutions. These observations suggest a role for TRBP in the mechanism of RISC formation as well as a means to improve siRNA functionality through modification of the RNA or its sequence.

The third focus was to develop a biocompatible and biodegradable cationic polymer capable of delivering siRNA into cells grown *in vitro*, with potential use later for *in vivo* applications. These nanoparticles (NPs) were generated with “click” chemistry, allowing for a systematic study of a range of variables contributing to siRNA binding. Strong siRNA binding required a NP containing a combination of primary and secondary amine groups to facilitate electrostatic interactions. Inclusion of alkyl chains enhanced binding, perhaps by causing vesicle-like formation, while polyethylene glycol (PEG) reduced overall binding affinity, consistent with its known charge shielding effects. The NPs were capable of delivering siRNAs to cells, but were unable to release them to initiate RNAi. Collectively, the work presented here provides a framework for identifying and engineering more active siRNA molecules by taking into account mRNA target structure and TRBP binding preferences, as well as achieving efficient cellular uptake of such engineered molecules with a novel type of polymeric nanoparticle.

Copyright by

JOSEPH A. GREDELL

2009

To My Family

ACKNOWLEDGEMENTS

I thank the many people who, in some form or another, contributed to the work described in this document.

First and foremost, I am grateful to my advisor, Professor Pat Walton, for his constant support and guidance. I am certain that his mentoring has made me into a better scientist, and person, that will serve me well in the future.

I also thank Professor Kris Chan, for closely assisting with my work over the years, in addition to Professors Mark Worden and Richard Schwartz, for serving as members of my committee; thank you for your critical, yet helpful, discussions, as well as for having taught several of the courses I enrolled in as part of my Ph.D. program. I am also thankful for the collaboration with Professor Greg Baker and Erin Vogel as we studied the polymeric nanoparticles that, although they "work" in many regards, are still a source of frustration.

Of course, I would be remiss if I didn't express my gratitude to the members of the Chan and Walton labs who I worked with on a daily basis and who provided support, critical analysis, and discussions. A special note of thanks goes out to Shireesh, Sri, Xuerui, Hemant, Shengnan, and Mike. In addition, I am grateful for the help of numerous undergraduate students who assisted with many of the daily experiments and were a pleasure to work with: Greeshma, Sophie, Dan, Mark, Jorge, Melissa, and in particular, Angela and Mike (Dittmer).

I acknowledge Michigan State University, the College of Engineering, the Department of Chemical Engineering and Materials Science, as well as the National Science Foundation and the National Institutes of Health for financial support.

I was fortunate to meet many other great people along the way, who helped me with my work, but more importantly were there as friends. In particular, Aaron, Brian, and Susan – thank you for listening, especially towards the end. And thank you to the MSU Men's Volleyball Club for showing me what a college experience really can be like.

Last but not least, I thank my mom, dad, and sister. There is too much to say, and too few words. But I would not be here today if not for your continuous love and support from the beginning.

Thank you.

TABLE OF CONTENTS

LIST OF TABLES.....	x
LIST OF FIGURES	xi
LIST OF ABBREVIATIONS	xiii
CHAPTER 1. INTRODUCTION.....	1
1.1 Significance	1
1.2 Background.....	2
1.3 RNA Interference mechanism	3
1.3.1 Human RNAi mechanism.....	8
1.3.2 dsRBPs involved in and related to RNAi	9
1.3.3 Non-specific effects - Immune response and off-target effects of siRNAs.....	14
1.4 Application of RNAi	15
1.4.1 Selection of highly active siRNA sequences.....	15
1.4.2 Specific uses of RNAi	19
1.4.3 siRNA delivery	22
1.4.4 Chemical modification of siRNA.....	27
1.5 Approach and specific aims.....	29
CHAPTER 2. IMPACT OF TARGET mRNA STRUCTURE ON siRNA SILENCING EFFICIENCY	32
2.1 Abstract.....	32
2.2 Introduction	34
2.3 Results	36
2.3.1 siRNA selection and mRNA target structure determination	36
2.3.2 Silencing activity of siRNAs targeting the EGFP mRNA.....	41
2.3.3 Consideration of larger siRNA dataset – data distribution.....	41
2.3.4 Distribution of base-pairs within the mRNA target and the effect on silencing activity	43
2.3.5 mRNA target loop size and location effects.....	46
2.3.6 siRNA guide strand structure effects.....	49
2.4 Discussion.....	52
CHAPTER 3 RECOGNITION OF siRNA BY TAR RNA BINDING PROTEIN (TRBP)	61
3.1 Abstract.....	61
3.2 Introduction	62
3.3 Results	64
3.3.1 TRBP recognition of small nucleic acids	65
3.3.2 TRBP asymmetry sensing in siRNAs.....	70
3.3.3 Altering siRNA asymmetry via mismatches	76
3.3.4 TRBP recognition of single-stranded small RNAs.....	82
3.4 Discussion.....	85

CHAPTER 4 POLYMERIC NANOPARTICLES FOR siRNA DELIVERY	91
4.1 Abstract.....	91
4.2 Introduction	93
4.3 Results and Discussion	95
CHAPTER 5 CONCLUSIONS AND FUTURE WORK.....	116
5.1 Conclusions	116
5.2 Future Work.....	117
5.2.1 siRNA selection algorithm	118
5.2.2 Further TRBP characterization.....	120
5.2.3 Characterization of additional RNAi-related proteins.....	121
5.2.4 Further nanoparticle investigations	123
APPENDICES	125
Materials and Methods for Ch. 2	125
Materials and Methods for Ch. 3	128
Materials and Methods for Ch. 4	137
BIBLIOGRAPHY	141

LIST OF TABLES

Table 1-1 siRNAs in clinical trials (http://clinicaltrials.gov).	21
Table 2-1 Details for EGFP targeting siRNAs experimentally validated in this study.	38
Table 2-2 Definition of RNA structures and siRNA activity within each group.	48
Table 2-3 Statistical analysis of siRNA activity between groups for mRNA target structures.....	50
Table 2-4 Statistical analysis of siRNA activity between groups for guide strand structures.....	51
Table 2-5 Distribution of siRNA functionality for each target structure classification. ...	53
Table 4-1 Binding strength (K) and cooperativity (n) of nanoparticles for siDNA as determined by gel shift assay.....	101
Table 6-1 Nucleic acids used in Ch. 3 (listed 5' to 3').	130
Table 6-2 Nucleic acid duplexes used in Ch. 3 (top strand listed 5' to 3', bottom strand listed 3' to 5').....	131
Table 6-3 Nucleic acid duplexes targeting pp-luciferase used in Ch. 3 (top strand listed 5' to 3', bottom strand listed 3' to 5').....	132
Table 6-4 Nucleic acid duplexes targeting sod1 used in Ch. 3 (top strand listed 5' to 3', bottom strand listed 3' to 5').....	133
Table 6-5 Nucleic acid duplexes targeting EGFP used in Ch. 3 (top strand listed 5' to 3', bottom strand listed 3' to 5').....	134

LIST OF FIGURES

Images in this dissertation are presented in color.

Figure 1-1 Mechanism of RNAi.....	4
Figure 1-2 siRNA structure.	6
Figure 1-3 Double stranded RNA binding domain structures of three dsRBPs.....	10
Figure 1-4 TAR RNA stem-loop secondary structure.....	12
Figure 2-1 mRNA predicted structures for siRNAs targeting EGFP.....	39
Figure 2-2 Local mRNA target structure groupings.....	40
Figure 2-3 mRNA target structure dependent gene silencing.	42
Figure 2-4 Distribution of siRNAs.	44
Figure 2-5 Influence of the number of base-pairs within the target site.....	45
Figure 2-6 Effect of window size on gene expression level.....	47
Figure 2-7 siRNA-mRNA interaction predictions.....	60
Figure 3-1 Characterization of recombinant TAR RNA binding protein (TRBP).	66
Figure 3-2 TRBP binding preferences for nucleic acids.	67
Figure 3-3 TRBP multimerization on siRNA.....	68
Figure 3-4 No MBP binding of small nucleic acids.	69
Figure 3-5 Nucleic acids used.	71
Figure 3-6 TRBP asymmetry sensing.....	73
Figure 3-7 Asymmetry sensing by human cell extract.	75
Figure 3-8 Computational analysis of first nucleotide influence on siRNA activity.	77
Figure 3-9 Statistical analysis for pair-wise comparison between first nucleotide on either end of the siRNAs.....	78
Figure 3-10 Mismatched siRNAs alter TRBP asymmetry.	80

Figure 3-11 TRBP recognition of short single-stranded RNAs.	83
Figure 4-1 Nanoparticle synthesis details.....	96
Figure 4.2 Visualization of nanoparticle delivery into cell culture by fluorescence microscopy.....	98
Figure 4-3 Nanoparticle complex formation with siDNA (analog of siRNA).	100
Figure 4-4 Nanoparticle functional groups added to the propargyl glycolide backbone.	104
Figure 4-5 Nanoparticle binding affinity relative to polymer backbone length.	107
Figure 4-6 Nanoparticle binding affinity relative to PEG length.	111
Figure 4-7 Nanoparticle binding affinity relative to alkyl chain length.	112
Figure 4-8 Nanoparticle size determined by Dynamic Light Scattering (DLS).....	113
Figure 4-9 Nanoparticle size and shape determined by Transmission Electron Microscopy (TEM).	114
Figure 4-10 Nanoparticle toxicity.....	115

LIST OF ABBREVIATIONS

Ago	Argonaute
AMD	age-related macular degeneration
AS	antisense strand
asODN	antisense oligodeoxynucleotide
ATP	adenosine triphosphate
BLAST	Basic Local Alignment Search Tool
BPEI	branched PEI
CHO	Chinese hamster ovarian
dFXR	drosophila fragile X protein
DLS	dynamic light scattering
DNA	deoxyribonucleic acid
DOTMA	N-[1-(2,3-dioleyloxy)propyl]-N,N,N-trimethylammonium chloride
dsDNA	double stranded DNA
dsRBD	dsRNA binding domain
dsRBM	dsRNA binding motif
dsRBP	dsRNA binding protein
dsRNA	double stranded RNA
EGFP	enhanced green fluorescent protein
eIF2 α	eukaryotic initiation factor 2 α
FAM	fluorescein amidite
FITC	fluorescein isothiocyanate
FPLC	fast performance liquid chromatography

HIV	human immunodeficiency virus
IFN	interferon
IL	interleukin
LF2k	Lipofectamine 2000
LNA	locked nucleic acid
LPEI	linear PEI
LTR	long terminal repeat
MBP	maltose binding protein
MDA	melanoma differentiation associated gene
Medipal	Merlin-Dicer-PACT liaison domain
MFE	minimum free energy
mRNA	messenger RNA
miRNA	microRNA
NCBI	National Center for Biotechnology Information
NP	nanoparticle
nt	nucleotide
PACT	PKR activator
PAZ	Piwi-Argonaute-Zwille
pDNA	plasmid DNA
PEI	polyethylenimine
PEG	polyethylene glycol
PIWI	P-element induced wimpy testis
PKR	protein kinase R

PLGA	poly(lactic-co-glycolic acid)
PLL	poly-l-lysine
PPGL	poly(propargyl glycolide)
pp-luc	Photinus pyralis (firefly) luciferase
Pre-miRNA	precursor miRNA
Pri-miRNA	primary miRNA
PTGS	post-transcriptional gene silencing
RDE	RNAi deficient
RHA	RNA helicase A
RIG	retinoic acid inducible gene
RISC	RNA induced silencing complex
RLC	RISC loading complex
RNA	ribonucleic acid
RNase	ribonuclease
RNAi	RNA interference
RNP	ribonucleoprotein
RSV	respiratory syncytial virus
shRNA	short-hairpin RNA
siDNA	small interfering DNA
siRNA	small interfering RNA
sod	Cu, Zn – superoxide dismutase
SS	sense strand
ssRNA	single-stranded RNA

ss-siRNA	single-stranded siRNA
STE	Salt-Tris-EDTA
TAR	trans-activating response element
Tat	transactivator of transcription
TBE	Tris-boric acid-EDTA
TEM	transmission electron microscopy
TLR	Toll-like receptor
TRBP	TAR RNA binding protein
Tudor-SN	Tudor staphylococcal nuclease
UTR	untranslated region
UV	ultraviolet
v/v	volume/volume
VEGF	vascular endothelial growth factor
VIG	vasa intronic gene

CHAPTER 1. INTRODUCTION

1.1 Significance

The flow of information from the DNA in our genomes to messenger RNA (mRNA), through the process of transcription, and the subsequent translation of protein from that mRNA, is a crucial process carried out in all the cells of our bodies. Disruption of that process can have severe effects, which can lead to disease. Understanding how our DNA is maintained and used for mRNA and protein production is therefore essential for developing methods for disease treatment and prevention. With the successful completion of the Human Genome Project in 2003, the DNA sequence for every putative human gene is known (Consortium, 2004; Venter et al., 2001). Current estimates place the number of protein-coding genes between 20,000-25,000 (Consortium, 2004). However, the precise function(s) of the proteins encoded by each gene is not yet known.

Accordingly, an extremely powerful research tool would be one by which gene expression can be specifically inhibited and the effect on cellular pathways observed. One tool currently being utilized for this purpose is RNA interference. RNA interference (RNAi) is a natural pathway that mediates sequence specific changes on mRNA levels through small interfering RNA (siRNA) molecules (Hannon, 2002). The resulting reduced protein levels are related directly to the Watson-Crick base-pairing between the siRNA and the complementary target mRNA. For scientists, RNAi is already proving useful, with synthetic siRNAs being used on a daily basis to target specific genes of interest. From the resulting changes in protein levels, protein function and interactions can in many instances be deduced. Additionally, these small RNA molecules are being

pursued as therapeutics, since they could potentially reduce the levels of proteins related to various diseases; some are even in clinical trials for treatment of diseases such as macular degeneration (<http://www.clinicaltrials.gov>).

The considerable investigation of the RNAi pathway since its discovery has led to an unfortunate conclusion, that not all siRNAs are able to reduce the level of their target gene. Selecting, testing, and characterizing siRNAs to find an effective sequence can be a time consuming and costly endeavor. Compound that with the fact that some siRNAs even have effects on the expression level of unintended mRNAs or cause immune responses (they resemble viruses in many ways), which can mask or preclude their desired effect, much remains to be learned to realize the full potential of RNAi-based applications. The work described here helps to improve the understanding of the mechanism of RNAi, provides additional means for selecting functional siRNA sequences, and presents a novel method for delivering siRNAs for therapeutic applications.

1.2 Background

In 1998, evidence was presented suggesting that dsRNA was responsible for suppression of gene expression in the nematode *Caenorhabditis elegans* (Fire et al., 1998). Double-stranded RNAs (dsRNAs) targeting well-characterized genes were injected into adult hermaphrodites, and changes in phenotype, such as body size and muscle formation, were observed. Small concentrations (only a few molecules) of dsRNA were required per cell to prevent expression of the target gene, and silencing endured for several cell divisions (Fire et al., 1998). Prior to this, several laboratories had observed a similar “co-suppression” effect in plants due to viral replication or transgene

expression (reviewed in (Baulcombe, 2004)). Today, these are classical examples of post-transcriptional gene silencing (PTGS) that fall under the category of RNA interference. This phenomenon results in potent and primarily specific gene silencing that occurs in most eukaryotes (Fire et al., 1998; Hannon, 2002) and is believed to exist as a mechanism for organisms to sustain genetic integrity by minimizing repetitive DNA and by preventing deleterious gene expression, such as expression of viral genes (Conklin, 2003). The profound effect that this pathway has had on our understanding of basic biology, particularly the role that RNA plays in cell regulation, as well as its potential for utilization as a research and therapeutic tool, led to the 2006 Nobel Prize in Physiology or Medicine being awarded to Andrew Fire and Craig Mello for its discovery.

1.3 RNA Interference mechanism

The mechanism of RNAi is composed of two distinct phases (Figure 1-1). In the initiation phase, long dsRNAs are cleaved by the ribonuclease (RNase) III enzyme Dicer into 21-27 nt non-coding RNA molecules called siRNAs (Bernstein et al., 2001; Elbashir et al., 2001a; Provost et al., 2002). They are then incorporated into the RNA induced silencing complex (RISC), which in the effector phase uses the siRNA as a template to bind to complementary mRNAs through Watson-Crick base-pairing. RISC cleaves the mRNA so that the corresponding protein cannot be translated; the effect is often referred to as gene silencing or knockdown.

There are several possible sources of dsRNAs greater than ~30 nucleotides (nt) in length that can be cleaved by Dicer to generate siRNAs. Some dsRNAs are naturally expressed (Watanabe et al., 2008), such as those produced by viruses or transposons (reviewed in (Saunders and Barber, 2003; van Rij and Berezikov, 2009)). Others can be

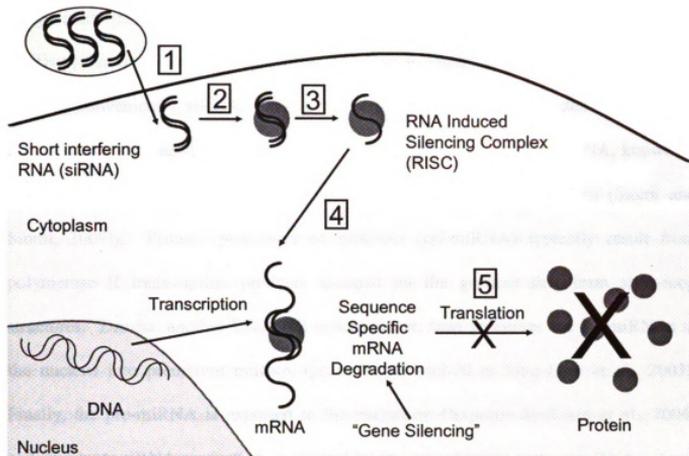


Figure 1-1 Mechanism of RNAi.

RNA interference is initiated by the successful delivery of siRNAs to the cell cytoplasm or by their endogenous expression (1). They are then recognized by the proteins comprising the RISC loading complex (RLC) (2) and a single strand of the siRNA duplex, the guide strand, is loaded onto RISC (3). That strand is then used to guide RISC binding to the target mRNA and subsequent cleavage (4), thereby preventing expression of the intended protein (5).

exogenously delivered via plasmids that, once transcribed, form short hairpin RNA (shRNA) that can be similarly processed by Dicer to yield siRNAs (reviewed in (Rossi, 2008)). Conveniently, siRNAs can also be chemically synthesized and delivered to cells, thus bypassing the need for Dicer processing. A related type of small RNA, known as microRNA (miRNA), is produced by a similar mechanism (reviewed in (Siomi and Siomi, 2009)). Primary precursors of miRNAs (pri-miRNA) typically result from polymerase II transcription products encoded by the genome that form stem-loop structures. Drosha, another RNase III endonuclease, then processes the pri-miRNAs in the nucleus into premature miRNA (pre-miRNA) ~65-70 nt long (Lee et al., 2003). Finally, the pre-miRNA is exported to the cytoplasm (Exportin-5) (Lund et al., 2004) and, similar to siRNA production, is cleaved by Dicer to yield the mature miRNA.

Much of the current understanding of how RNAi works is based upon studies done using *Drosophila melanogaster*. In the fruit fly, the second of two Dicer isoforms (Dicer-2) is responsible for processing long dsRNAs into siRNAs (Bernstein et al., 2001). Due to the orientation and processing of the RNA-binding and RNase domains of Dicer-2, the resulting siRNAs are composed of two complementary strands having 5'-phosphate groups and 2 nt overhangs on either 3'-end (Bernstein et al., 2001) (Figure 1-2). These two distinct structural features are critical for their downstream activity yet are independent of their nucleotide sequence. Dicer-2 and an associated double-stranded RNA binding protein (dsRBP), R2D2, then jointly interact with the siRNA (Liu et al., 2003). The heterodimer of Dicer-2/R2D2 positions itself on the siRNA so that R2D2 binds to the more stable end of the duplex (Tomari et al., 2004) while the PAZ domain of Dicer-2 recognizes the 3'-OH group on the opposing 2 nt overhang (Ma et al., 2004). The

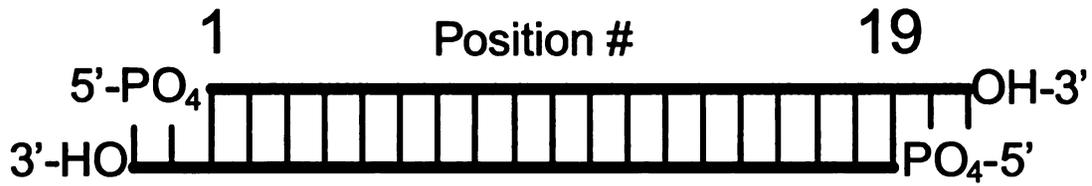


Figure 1-2 siRNA structure.

siRNAs are composed of two ~21 nt long complementary RNA strands held together by Watson-Crick base-pairing. The strands are slightly offset to have 2 nt overhangs on either 3'-end and each 5'-end is phosphorylated. The top strand here is designated as the "antisense" or "guide" strand and is complementary to the target mRNA. The bottom strand is referred to as the "sense" or "passenger" strand.

positioning of these three molecules, known as the RISC loading complex (RLC), occurs based on the thermodynamics of the siRNA and the 5'-phosphorylation status (Tomari et al., 2004). The strand bound by Dicer-2 at the 5'-end is used to guide RISC to the target mRNA and is therefore known as the guide strand. Next, the endonuclease Argonaute (Ago) 2 associates with the RLC and cleaves the passenger strand in an ATP-dependent step (Matranga et al., 2005). The passenger strand fragments are released for further degradation (Liu et al., 2004) and only the guide strand is preferentially retained in active RISC (Matranga et al., 2005; Tomari et al., 2004).

Active RISC subsequently uses the selectively incorporated guide strand as a template to bind mRNAs containing regions of Watson-Crick base-pairing complementarity. It is believed that the first 2-8 nts of the guide strand, known as the seed region, primarily contribute to target recognition by providing the necessary binding energy (Haley and Zamore, 2004). Perfect binding between guide strand-loaded RISC and mRNAs, such as tends to be the case when using an siRNA, leads to Ago2-mediated cleavage of the cognate mRNA between bases 10 and 11 of the siRNA/mRNA duplex relative to the 5'-end of the guide strand (Elbashir et al., 2001b; Rand et al., 2005; Song et al., 2004). However, mismatches between the guide strand and the mRNA, often encountered when RISC is loaded with miRNAs, frequently can result in translational repression (Doench and Sharp, 2004). For both mechanisms, translation of the corresponding protein is inhibited, resulting in the so-called gene silencing that is the hallmark of RNAi. Several additional proteins can associate with active RISC, such as Tudor staphylococcal nuclease (Tudor-SN), vasa intronic gene (VIG), and the *Drosophila* homolog of fragile X-related gene (dFXR) (Caudy and Hannon, 2004; Caudy et al., 2003;

Caudy et al., 2002; Ishizuka et al., 2002). However, their role in RNAi is unknown and appears to be nonessential as only Dicer-2/R2D2/Ago2 is necessary for function.

1.3.1 Human RNAi mechanism

The RNAi mechanism is largely conserved among eukaryotes. Recently, numerous studies have elucidated many of the details for human RNAi. Human Dicer, ~200kDa, can bind and cleave dsRNAs into siRNAs (MacRae et al., 2007; MacRae et al., 2006; Pellino et al., 2005), as in *Drosophila*, as well as bind short single-stranded siRNAs (ss-siRNAs) (Kini and Walton, 2007; Lima et al., 2009). In the cases where siRNAs are exogenously delivered, if either 5'-end lacks a phosphate, as with chemically-synthesized siRNAs, a phosphate is added by Clp (Weitzer and Martinez, 2007). The siRNAs then become part of the RLC with Dicer and the ~35kDa TAR RNA binding protein (TRBP) (Chendrimada et al., 2005; Gregory et al., 2005; Haase et al., 2005; MacRae et al., 2008). TRBP is a dsRBP with homology to R2D2 and is hypothesized to perform a similar role in the RLC (Forstemann et al., 2005; Haase et al., 2005; Murphy et al., 2008). Once Dicer/TRBP associate with an siRNA, one of four Argonaute proteins joins the complex (reviewed in (Hutvagner and Simard, 2008; Siomi and Siomi, 2009)). Only RISC formed with Ago2 is capable of cleaving target mRNAs (MacRae et al., 2008; Meister et al., 2004; Rivas et al., 2005; Vickers et al., 2007). Interestingly, recombinant Ago2 can bind ss-siRNAs alone and use them to guide target mRNA cleavage (Rivas et al., 2005). As with *Drosophila*, several additional proteins can associate with RISC, most likely through specific interactions with a particular RISC component. For instance, Dicer is known to interact directly with TRBP and the Protein Kinase R (PKR) activator (PACT) (discussed more below) (Kok et al., 2007; Laraki et al., 2008). Similarly, Ago2 associates with the

DEAD-box helicase DP103 (Gemin3 and ddx20) and P-body components, among others (Donker et al., 2007; Hutvagner and Zamore, 2002; Jakymiw et al., 2005; Lian et al., 2007; Liu et al., 2005a; Liu et al., 2005b; Meister et al., 2005).

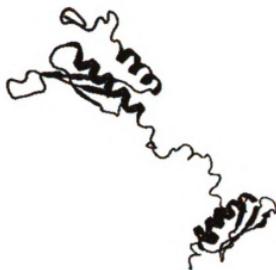
1.3.2 dsRBPs involved in and related to RNAi

Double stranded RNA binding proteins (dsRBPs) contribute to numerous critical cellular pathways, including gene silencing, RNA editing, protein phosphorylation, transcription, and translation (Chang and Ramos, 2005). By definition, they bind to dsRNA, typically doing so through dsRNA binding domains (dsRBDs). dsRBDs maintain a highly-conserved $\alpha\beta\beta\alpha$ (Figure 1-3) structure after folding of ~65-68 amino acids that can interact with A-form RNA via the 2'-OH and backbone phosphate groups (Doyle and Jantsch, 2002; Nanduri et al., 1998; Ramos et al., 2000; Ryter and Schultz, 1998; Saunders and Barber, 2003). Additionally, most dsRBDs appear to rely on at least 16-20 base-pairs of dsRNA for recognition (Manche et al., 1992; Ryter and Schultz, 1998; Zheng and Bevilacqua, 2004) since shorter sequences may not allow sufficient protein/RNA contacts, as illustrated by the crystal structure of a dsRBD from *Xenopus* Xlrbpa with a short non-physiological dsRNA comparable in size to an siRNA (Ryter and Schultz, 1998). However, two reports show that as few as 11 base-pairs (equivalent to one helical turn of dsRNA) can be sufficient for binding by a single dsRBD (Manche et al., 1992; Nanduri et al., 1998). Binding is generally acknowledged to be sequence independent because the major groove, where specific base interactions are mediated, is narrow and deep and therefore inaccessible (Chang and Ramos, 2005; Fierro-Monti and Mathews, 2000; Seeman et al., 1976). However, tolerance for disruptions of the dsRNA

TRBP



PKR



PACT



Figure 1-3 Double stranded RNA binding domain structures of three dsRBPs.

The solution structures for the dsRBD of TRBP (dsRBD 2; Protein Data Bank# 2cpn), PKR (both dsRBDs; #1qu6), and PACT (#2dix) reveal conserved $\alpha\beta\beta\alpha$ structures that are believed to mediate contacts with the 2'-OH and phosphate groups of the dsRNA (reviewed in (Chang and Ramos, 2005)).

helix exist and may in some instances even contribute to binding (Bevilacqua et al., 1998; Nallagatla et al., 2007). Furthermore, dsRBPs do not tend to bind ssRNA unless they form secondary structures like hairpin loops (Saunders and Barber, 2003).

A classic source of dsRNA is viral infection, and accordingly, it is no surprise that many dsRBPs participate in the innate immune response. More recently noteworthy for its presence in human RISC, TRBP was originally identified by its ability to bind the Trans-activating region (TAR) located in the 5'- and 3'-long terminal repeats (LTRs) of all human immunodeficiency virus (HIV-1) mRNA transcripts (Gatignol et al., 1991). TAR RNA forms a classic stem-loop secondary structure (Figure 1-4) where the stem, bulge (nts 23-25), and loop (nts 30-35) are important for transactivator of transcription (Tat) binding and subsequent mRNA transcription (Berkhout and Jeang, 1991; Rana and Jeang, 1999). TRBP facilitates Tat binding, possibly by altering the local structure of TAR or by mediating binding of other cofactors, and enhances transcription rates, aiding HIV-1 infection (Erard et al., 1998; Gatignol et al., 1991). TRBP has since been shown to have oncoprotein-like behavior as its over-expression promotes proliferation (Benkirane et al., 1997; Eckmann and Jantsch, 1997). Therefore, TRBP provides a direct link among several biological pathways, including viral recognition and gene silencing, by the manner in which it interacts with various types of dsRNA.

Interestingly, a related dsRBP, Protein Kinase R (PKR), can also bind TAR RNA (with a dissociation constant of ~ 1 -100 nM), but, rather than facilitating HIV-1 expression, it suppresses function (Bevilacqua and Cech, 1996; Kim et al., 2006). It was therefore not surprising to find that TRBP inhibits the antiviral effect of PKR (Benkirane et al., 1997). PKR binds dsRNA via its two dsRBDs that work in tandem

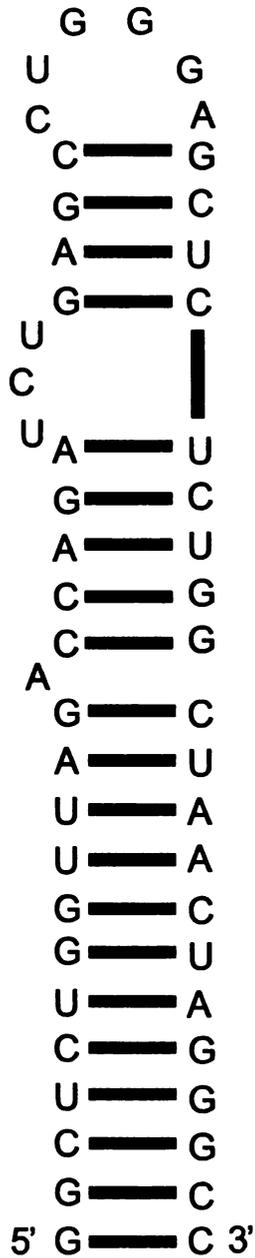


Figure 1-4 TAR RNA stem-loop secondary structure.

TRBP was discovered due to its binding of this TAR RNA stem-loop structure that forms at the 5'- and 3'-ends of HIV-1 transcripts.

(Bevilacqua and Cech, 1996; Bevilacqua et al., 1998; Carlson et al., 2003). Upon binding, PKR homodimerizes and activates by autophosphorylation, initiating a signaling cascade that culminates in cytokine production (specifically the interferons, IFN- α and IFN- β) and a generalized inhibition of protein translation (through eIF2 α) (Cosentino et al., 1995; Garcia et al., 2006; Patel et al., 1995). These two results contribute to the innate immune response to viral infection by killing the infected cell in an attempt to prevent spread of the infection.

It is not quite clear how TRBP and PKR have opposing roles in their response to dsRNA despite similar means of binding. One theory is that TRBP sequesters the dsRNA to preclude recognition by PKR. It is also possible that TRBP and PKR bind simultaneously to a dsRNA (Cosentino et al., 1995). Like PKR, TRBP contains two dsRBDs at its N-terminus that are responsible for dsRNA binding, and its affinity for TAR RNA and other dsRNA appears to be comparable to that observed for dsRNA binding by PKR, ~1-100 nM (Bevilacqua and Cech, 1996). However, TRBP only appears to require the second dsRBD for high affinity binding (Daviet et al., 2000). Also interesting is that TRBP binding to dsRNA appears to be uncooperative; that is, it does not show increased affinity for longer sequences (Parker et al., 2008). This is in contrast to PKR and RDE-4 (a dsRBP from *C. elegans* that participates in siRNA production) which show a marked difference in affinity for longer dsRNAs (Parker et al., 2008). TRBP and PKR can also heterodimerize directly through the C-terminal domain of TRBP, called the Medipal domain (Daher et al., 2001), as well as heterodimerize with a third dsRBP, PACT (Laraki et al., 2008). PACT helps to activate PKR (Chendrimada et al., 2005; Haase et al., 2005; Kok et al., 2007; Li et al., 2006), but, because it shares

~42% homology with TRBP, it is hypothesized to assist in siRNA production (Haase et al., 2005; Kok et al., 2007). Similarity among the dsRBDs of these (Figure 1-3), and a number of other proteins involved directly in RNAi, illustrates the likely importance of these domains and their interactions with RNAs in achieving silencing.

1.3.3 Non-specific effects - Immune response and off-target effects of siRNAs

Given the similarity of TRBP, PKR, and PACT as dsRBPs and their role in recognition of dsRNA for RNAi or during viral infection, it is not surprising that siRNAs can also stimulate an immune response (Bridge et al., 2003; Hornung et al., 2005; Judge et al., 2005; Kariko et al., 2004; Kim et al., 2004; Sledz et al., 2003). Other cytoplasmic proteins, such as the RNA helicases RIG-1 and MDA-5, contribute to this effect, by activating pathways that stimulate the interferons, like PKR does, and other inflammatory cytokines (Yoneyama et al., 2004). While each of these proteins is predominantly localized within the cytoplasm, the immunostimulatory effect of siRNAs is more often associated with the transmembrane Toll-like Receptors (TLRs). The family of ~10 TLRs in humans is traditionally responsive to viruses or bacteria components, with each receptor showing specificity towards a particular type of pathogen (Meylan and Tschopp, 2006). TLR3 is documented for recognizing dsRNA independent of sequence, including siRNA in some instances (Kariko et al., 2004), while TLR7/8 is more often associated with binding ssRNAs (Kleinman et al., 2008; Sioud, 2006). However, TLR7/8 have recently been implicated in initiating an immune response to certain siRNA sequence motifs (Hornung et al., 2005). These TLRs (3/7/8) are not highly expressed on most cell

lines and that could explain why early studies with siRNAs in human cell culture did not detect an immune response (reviewed in (Judge and Maclachlan, 2008)).

In addition to the immune response occasionally stimulated by siRNAs, siRNAs can have unanticipated effects on many other untargeted genes by being only partially complementary to mRNAs besides the intended target (Fedorov et al., 2006; Jackson et al., 2003; Scacheri et al., 2004). It appears that target recognition is mediated through similarity with the “seed” region of the siRNA, the first 2-8 bases, or with the passenger strand (Jackson et al., 2006; Lai, 2002).

1.4 Application of RNAi

The primary goal of applying RNAi for research or therapeutic purposes, like any other drug, is to achieve targeted and controllable changes in the cell while minimizing undesirable effects. Much of the difficulty surrounding the application of siRNAs is due to the variability in silencing activity of different sequences, where, unfortunately, the proportion of siRNAs that are successful in repressing their target gene, termed active or functional siRNAs, is low, not unlike what was found for antisense oligonucleotides (asODNs), another technique for inhibition of gene expression by targeting the mRNA (Stein, 1998; Vickers et al., 2003). Thus, a common goal of RNAi-related research is to develop siRNAs that maximize reduction of the target gene at the lowest possible concentration while avoiding non-specific effects.

1.4.1 Selection of highly active siRNA sequences

The identification of highly active siRNA sequences is complicated by the sheer size of the human genome. This makes a randomized approach difficult, since early work

with RNAi found that randomly selected siRNA activity could vary greatly (Amarzguioui and Prydz, 2004). Initial strategies for improving the selection focused on eliminating sequences based on high (>70%) or low (<40%) GC content and stretches of greater than four consecutive identical bases (e.g., GGGG) (Elbashir et al., 2002). These two guidelines made siRNA synthesis more efficient but were not based on mechanistic understanding.

Since then, as the number of siRNAs tested has increased, empirical rules have been proposed that can be used as part of elaborate computational algorithms that use tens of parameters to predict which sequences are active and even what their activity will be (Ge et al., 2005; Lu and Mathews, 2007; Shah et al., 2007; Shao et al., 2007; Vert et al., 2006). In many cases, these rules merely result from statistical sampling of large numbers of sequences and have limited mechanistic implications. More elaborate selection algorithms have been developed that further discriminate the most important siRNA structural and sequence features.

Perhaps the most important rule in siRNA design, duplex asymmetry, is derived from the mechanistic understanding that RISC can sense the relative end stabilities of the siRNA duplex. In theory, both strands of the siRNA may be incorporated into active RISC, with approximately half of RISCs able to silence the target complementary to the guide strand and the other half silencing any targets that happen to be complementary to the passenger strand. However, the strand whose 5'-end is less stably hybridized within the siRNA duplex becomes preferentially incorporated into active RISC (Khvorova et al., 2003; Schwarz et al., 2003; Tomari et al., 2004). The result is that a higher proportion of active RISCs contain the guide strand (for the desired target), leading to more active

target silencing. Many of the positional base preferences that have been identified and are implemented in siRNA design algorithms tend to yield the desired differential stability between the two ends (Jagla et al., 2005; Lu and Mathews, 2007; Reynolds et al., 2004; Shao et al., 2007; Ui-Tei et al., 2004).

Although duplex asymmetry can lead to favorable guide strand selection after passenger strand cleavage by Ago2, it is possible that the retained guide strand forms secondary structure that can inhibit Ago2-loading and subsequent target recognition (Patzel et al., 2005). Rather than experimentally determine guide strand structure, it is more commonly predicted using separately developed algorithms (discussed in more detail below). Strands that form little to no structure are then weighted favorably in the siRNA selection algorithm (Koberle et al., 2006; Patzel et al., 2005). However, the impact of guide strand structure on selecting effective siRNAs is still under debate and is not incorporated into all current algorithms (Lu and Mathews, 2007).

Once the guide strand is loaded onto Ago2, RISC must locate and bind to complementary mRNAs. As one might expect, the accessibility of the target mRNA can have considerable influence on the accompanying level of silencing, as had been found for asODNs (Vickers et al., 2000; Walton et al., 2002). mRNA transcripts within the cell are known to exist in complexes with numerous ribonucleoproteins (RNPs) that are important for stabilizing the RNA and for protein synthesis. Naturally, these RNPs pose a hindrance for RISC binding, and vice versa, but not in a way that is at present predictable. However, accessibility can also be limited by secondary structures formed by the mRNA. Whereas siRNA and miRNA sequences are only ~21 nt long, mRNAs can be several 1000 nt long and therefore tend to possess significant numbers of

intramolecular base-pairs (e.g., the mRNA transcript encoded by the enhanced green fluorescent protein (EGFP) plasmid is predicted to have more than 50% of the nucleotides involved in base-pairs (Gredell et al., 2008)). Several programs exist that predict RNA structures using “nearest-neighbor” free energy values experimentally determined for short RNA sequences (Mathews et al., 1999). One class of program (e.g., mfold/UNAFold) minimizes the global free energy of the entire RNA sequence by extrapolating from the nearest-neighbor values (Markham and Zuker, 2008; Zuker, 2003). A second class of program (e.g., Sfold, RNA Vienna Package) utilizes a statistical partition function approach based on the Boltzmann ensemble of all allowable secondary structures, from which the most likely structure is selected (Ding et al., 2004; Hofacker et al., 1994). The programs quickly calculate siRNA guide strand structure (21 nt long), but the calculation time and memory requirements scale as the sequence length cubed, so mRNAs can require considerably more time and computational capacity (Lu and Mathews, 2007).

Several reports have used these programs to predict mRNA structures and demonstrate a direct impact on siRNA efficacy (Ameres et al., 2007; Bohula et al., 2003; Brown et al., 2005; Far and Sczakiel, 2003; Lu and Mathews, 2007; Overhoff et al., 2005; Schubert et al., 2005; Shao et al., 2007; Vickers et al., 2003; Westerhout and Berkhout, 2007; Yoshinari et al., 2004). Gene silencing decreased when the orientation or the degree of partial base-pairing of a target construct was varied for a single siRNA (Schubert et al., 2005; Westerhout and Berkhout, 2007). This idea was extended to full-length transcripts (ICAM-1 and *survivin*) where siRNAs targeting inaccessible regions (i.e., those with extensive secondary structure) were ten-fold less active than accessible

sites (Overhoff et al., 2005). One particularly intriguing indication of the importance of target mRNA structure on RNAi is illustrated by the ability of HIV-1 to overcome silencing by mutating the siRNA target sequence, doing so in a fashion that also alters its local RNA secondary structure (Leonard et al., 2008; Westerhout et al., 2005).

Today, the best siRNA selection algorithms rely most heavily on rules that support incorporation of the appropriate guide strand, based on siRNA asymmetry, and account for mRNA target structure effects (Lu and Mathews, 2007; Shao et al., 2007). When trained with large data sets consisting of results from experiments with hundreds of different siRNAs, they can predict siRNA sequences for new targets with upwards of 90% accuracy (Lu and Mathews, 2007; Shao et al., 2007). Furthermore, attempts are being made to select against sequences resulting in non-specific effects, including off-target silencing and immune responses (Fedorov et al., 2006; Jackson et al., 2003; Reynolds et al., 2006; Scacheri et al., 2004; Sledz et al., 2003). Considerable effort is also being made by privately held companies to predict siRNAs. It is unknown what the various proprietary algorithms incorporate, but they are clearly capable of selecting effective siRNAs and almost certainly include many of the variables discussed above. Companies are also beginning to offer sets of validated siRNAs for human, mouse, and rat genomes that have been identified and experimentally validated to possess "high activity" by some definition.

1.4.2 Specific uses of RNAi

The most common application of RNAi is for studying gene and protein function. In this situation, one or more genes can be targeted for silencing by one or more siRNAs. The effect on the cell can then be analyzed to deduce function of the missing protein.

Results such as these are published virtually every day and have already provided much insight into biological function.

However, it is becoming of greater interest to use small RNAs to target genes of known function to improve the expression of recombinant proteins in plants and modified mammalian cells (e.g., Chinese hamster ovary (CHO)). Currently, 60-70% of all recombinant protein pharmaceuticals are produced in mammalian cells (Wurm, 2004). These proteins tend to be of high value and therefore slight gains in output can have vast economic benefits. Rather than use siRNAs to initiate RNAi, which result in transient changes to gene expression, the host genome can be modified to stably produce shRNAs that permanently alter expression of genes/pathways detrimental to protein production (Cox et al., 2006; Dodo et al., 2008). The targeted genes tend to lead to increased cell density and growth rates, and improved protein solubility, stability, and membrane permeability (Allen et al., 2008; Allen et al., 2004; Kim et al., 2007; Tang et al., 2007; Zhu and Galili, 2004).

A valuable use of the RNAi pathway is to treat human disease caused by unregulated protein expression. siRNAs are already being tested *in vivo*, and several are entering, or are currently in, clinical trials (<http://clinicaltrials.gov>). These first generation therapeutic siRNAs, summarized in Table 1-1, are typically delivered locally by physical means and target well-studied pathways. Furthermore, many of the siRNAs are chemically modified to enhance their stability *in vivo* (discussed below). In some cases, the method of delivery (discussed below) is as much the focal point of study as the response to the siRNA itself.

Table 1-1 siRNAs in clinical trials (<http://clinicaltrials.gov>).

siRNA	Source	Condition	Target	Delivery Method	Phase
Bevasiranib (Cand5)	OPKO Health	Wet Age-Related Macular Degeneration	VEGF	Intravitreal	III
AGN211745 (Sirna-027)	Allergan/Merck	Diabetic Macular Edema	VEGF	Intravitreal	II
PF-04523655	Quark/Pfizer	Wet Age-Related Macular Degeneration	VEGFR1	Intravitreal	II
(REDD14NP)		Wet Age-Related Macular Degeneration	DDIT4	Intravitreal	II
ALN-RSV01	Alyniam	Respiratory Syncytial Virus Infection	RSV-N	Intranasal	II
I5NP (Aki-5)	Quark	Acute Renal Failure	TP53	Intravenous, saline	I
CALAA-01	Calando	Solid Tumor	RRM2	Intravenous, polyplex	I
TD101	Transderm	Pachyonychia Congenita	KRT6A	Subcutaneous, saline	I

In each of these clinical trials using an siRNA as the therapeutic agent, the intention is to reduce the expression of a specific gene. As with any therapeutic, the expected mechanism of action may not always be the actual mechanism. Recently it was found that an siRNA targeting the vascular endothelial growth factor (VEGF) or its receptor (for treatment of choroidal neovascularization due to wet amyloid macular degeneration - AMD) could achieve an effect on the target gene by inadvertently activating TLR3 (Kleinman et al., 2008). The resulting immune stimulation, as characterized by the induction of IFN- γ and IL-12, was principally responsible for the therapeutic effect, independent of the RNAi pathway. In two other studies, immune stimulation due to an shRNA was linked to excess guide strand levels, perhaps from TLR7/8 responsiveness, or from competition with endogenous miRNAs that caused a saturation of some of the RNAi pathway proteins, particularly Exportin-5 (Grimm et al., 2006; McBride et al., 2008). These results from shRNA mediated silencing further support investigation of the use of siRNA-based RNAi, as siRNAs are not trafficked and processed in the same manner as shRNAs and miRNAs. Nonetheless, understanding strategies for mitigating the side-effects of stable *in vivo* silencing will be useful for development of treatments for chronic diseases as well as for establishing cell lines with constitutively modified gene expression.

1.4.3 siRNA delivery

While the improvement of algorithms that effectively identify highly functional siRNAs is a point of emphasis for *in vitro* applications of RNAi, the main limitation for *in vivo* work, and consequently usage in a human clinical setting, is delivery. In some lower organisms, such as worms, dsRNA (siRNA) can be eaten, absorbed, or injected,

resulting in highly efficient, systemic delivery (Fire et al., 1998; Hannon, 2002). However, in mammals, systemic uptake of naked siRNA (siRNA alone with no delivery agent) is typically poor, as nucleic acids do not freely diffuse across the cellular membrane. This effect is greatly compounded due to the rapid degradation of siRNAs by nucleases (< 1hr half-life in serum), binding by other factors found in the blood, and by clearance by the kidneys (Zhang et al., 2007). Although viral methods have the ability to overcome these limitations by incorporation of silencing constructs into the infected cells (reviewed in (Manjunath et al., 2009)), viral gene therapy methods still suffer from a number of safety issues that have caused multiple gene therapy clinical trials to be abandoned or curtailed. Thus, much effort is being spent developing non-viral delivery methods for siRNAs so that their therapeutic potential can be realized.

In some instances, uptake can be forced through physical methods, including intravascular injection, ultrasound, electroporation, and gene guns (Wolff and Rozema, 2008). In other circumstances, delivery can be accomplished by aerosolizing the siRNA for administration to the lungs, or by application of siRNA in a topical cream for dermal delivery (de Fougerolles et al., 2007; Wang et al., 2008). The siRNAs currently in clinical trials are delivered either by injection into the eye (for age-related macular degeneration, AMD) or inhalation to the lungs (for respiratory syncytial virus, RSV) (de Fougerolles, 2008). Unfortunately, these methods are not viable for delivery to deep tissues or tumors, targets of significant interest for siRNA therapeutics. As an alternative, both lipid and polymer based reagents are being pursued as delivery vehicles that can be applied systemically, with the potential for localized targeting, and that may even allow

transport through the blood brain barrier for treatment of neurological disorders (Pardridge, 2007).

Due to their phosphate backbone, all nucleic acids are negatively charged. Accordingly, delivery vehicles typically utilize electrostatic interactions between the nucleic acid backbone and either cationic lipids (lipoplexes) or cationic polymers (polyplexes). Condensing the siRNA in the vehicle can be achieved through simple mixing of the siRNA and vehicle in an aqueous buffer and allowing them to self-assemble or by more complex techniques such as spray drying (Takashima et al., 2007). The resulting complexes, which can range in size from ~50-300 nm, have been applied *in vitro* or *in vivo*. In several of the conditions studied to date, the complex charge tends to be slightly positive overall, which facilitates contact with the negative charge of the cellular membrane without causing substantial interactions with blood components that would lead to complement activation (Bartlett and Davis, 2007a, b). For systemic delivery, while there is no universally optimal size, it is generally considered that complexes must be large enough to protect the siRNA from nuclease digestion and renal clearance (> 10nm) but also small enough (< 70-100nm) to allow access to cells such as hepatocytes through capillary circulation (Bartlett and Davis, 2007b; Wolff and Rozema, 2008). After association with the extracellular membrane, the complexes are then endocytosed (Zuhorn et al., 2007). The siRNAs must then escape the endocytic vesicles into the cytosol to initiate RNAi.

The traditional structure of lipid-based delivery reagents is that of a fatty acid with a cationic head group and a non-polar hydrocarbon tail. Use of this type of transfection was first demonstrated using DOTMA with plasmids (Felgner et al., 1987).

Early RNAi transfection was performed using lipids developed for delivery of either plasmids or antisense oligonucleotides. Since then, a variety of proprietary lipid formulations have been developed that are specifically designed for siRNA transfection to cultured cells, such as Lipofectamine RNAiMAX (Invitrogen) or siPORT NeoFX (Applied Biosystems - Ambion). These reagents can be very effective at delivering siRNAs to diverse cell lines, but, unfortunately, they can be toxic at concentrations only moderately higher than concentrations required for effective siRNA delivery; *in vivo* experiments have also demonstrated similar toxicity concerns (Zhang et al., 2007). While generally not an issue for cultured cell applications, this small therapeutic window limits the prospects for use of these types of delivery agents for clinical applications. Recently, a combinatorial approach was used to test lipid-like compounds termed “lipidoids” to deliver siRNAs, finding that a diverse subset caused extensive silencing in conditions ranging from cell culture to non-human primate animals (Akinc et al., 2008). In many cases, these molecules were also found to have acceptable toxicity profiles *in vivo*. These readily synthesized materials provide an alternative to traditional lipids, expand the set of available transfection reagents, and provide insight as to the physical and chemical characteristics necessary for the most effective lipid vehicles. However, due to the relatively limited chemical and structural diversity in lipid and lipid-like vehicles, considerably more effort at creating *in vivo* delivery vehicles has been invested in using polymeric vehicles for nucleic acid delivery, in general, and for delivery of siRNAs, specifically.

Polymeric vehicles have strong potential for siRNA delivery applications because they provide protection from nucleases and other serum components and facilitate

endocytosis of the siRNA and its release into the cytosol. Of these, polyethylenimine (PEI) is perhaps the most commonly used and well-studied (Kircheis et al., 2001). At physiologic pH, the amine groups of PEI are protonated, providing the positive charge necessary to condense the nucleic acid (Clamme et al., 2003). Both linear PEI (LPEI) and branched PEI (BPEI) have shown encouraging results in silencing applications, with some molecular weights also demonstrating delivery of active siRNAs *in vivo* (Urban-Klein et al., 2005). In several instances, modifications, such as biomolecule addition (e.g., cholesterol, antibodies/peptides, or aptamers) or crosslinking have been made to minimize the toxicity associated with the larger molecular weight PEIs and to improve and target delivery (Swami et al., 2007). A particularly common modification is the addition of polyethylene glycol (PEG), which partially shields the charges on the polymer and siRNA, thereby improving membrane interactions, increasing product release, and decreasing toxicity (Wolff and Rozema, 2008). Interestingly, there appears to be an optimal amount and length of PEG per polymer, although the combinations studied have not led to a consensus set of guidelines (Brus et al., 2004; Kunath et al., 2002; Mao et al., 2006).

Although PEI is perhaps the most commonly studied polymer, two other considerably more complex systems have shown intriguing *in vivo* results. The first is called a “Dynamic PolyConjugate” (Rozema et al., 2003; Rozema et al., 2007; Wakefield et al., 2005). This vehicle contains multiple functional entities to address delivery in a stepwise manner. In a manner akin to a multi-stage rocket, groups are released from the conjugate once they have served their purpose, thus preventing interruption of downstream events. Starting with a membrane-active poly butyl amino vinyl ether

(PBAVE) backbone, PEG and N-acetylgalactosamine (for hepatocyte targeting) were attached via maleamate linkages. The siRNA is also covalently linked to the backbone via a disulfide bond. The maleamate linkages are reduced in the endosomal vesicles, exposing the PBAVE, which lyses the endosomal vesicles. The PBAVE-siRNA disulfide bond is reduced once the complex enters the cytoplasm, thereby protecting the siRNA from degradation until it is released in the compartment where it can access the RNAi machinery. Two different siRNAs delivered using this vehicle silenced their target genes by ~60-80% (Rozema et al., 2007).

A second system is built by self-assembly of an siRNA with a cyclodextrin-modified PEI, PEG, and transferrin (for targeting to tumor cells) (Hu-Lieskovan et al., 2005; Pun et al., 2004). This complex shows low toxicity and effective silencing of the target gene in a mouse model and in non-human primates. Together, these two polymeric systems highlight the utility of incorporating factors for targeted delivery and endosomal release. However, they also illustrate the potential complexity of the polymers required to address the many limitations that restrict delivery of active siRNAs *in vivo*.

1.4.4 Chemical modification of siRNA

The numerous studies utilizing siRNA underscore the complexity of the various overlapping pathways that can be initiated by dsRNA and that care must be taken when eventually progressing to *in vivo* applications. Chemical modification of the siRNA is one avenue of pursuit to address the problems of nuclease degradation, immune activation, off-target effects, cell uptake, and pharmacokinetics, essentially aiming to increase the longevity and specificity of the siRNA in the cellular environment.

Coincidentally, many of these problems were previously encountered with asODNs and thus those results have at least partially guided siRNA development in this regard.

There are limited ways in which the siRNA can be modified. The overall shape (i.e., overhangs) or length can be changed, where siRNAs only a few nucleotides longer than 21 base-pairs can sometimes be processed efficiently by Dicer and may facilitate guide strand loading into RISC (Kim et al., 2005; Rose et al., 2005; Siolas et al., 2005). Alternatively, chemical modifications can be made to the three key regions of each nucleotide: the phosphodiester backbone, the ribose ring, or the nucleoside base. Modifications to the ribose sugar, specifically at the 2'-position, are probably the most common alterations used for siRNAs. The groups frequently added, such as 2'-O-methyl, 2'-fluoro, and locked nucleic acids (LNA) (Corey, 2007) interfere with hydrolysis. In some cases, they even enhance activity relative to unmodified sequences (Elmen et al., 2005; Terrazas and Kool, 2009). However, bulkier groups at the 2' position are less well tolerated and start to have a detrimental effect on silencing potency. A variety of backbone modifications are available, with phosphorothioate (PS) linkages being routinely used because they specifically enhance the resistance of the backbone to cleavage by RNases. Unfortunately, substantial PS modifications result in increased cytotoxicity (Corey, 2007; Rana, 2007). Boranophosphate (BO) linkages, while having been studied less frequently and being limited in scale by synthesis techniques, appear to offer similar benefits (Corey, 2007). Base modifications are more limited, but can also improve resistance to nuclease digestion. However, that tends to come at the price of silencing because many of the modifications reduce the hydrogen bonding and therefore decrease hybridization affinity. Other modifications that were originally devised to

improve the biodistribution of antisense oligonucleotides, such as direct conjugation to cholesterol, receptor ligands, and transport peptides, can potentially be applied for siRNAs as well, provided they do not prevent recognition of the modified siRNA by RNAi proteins (reviewed in (de Fougerolles et al., 2007)).

Two more recent reports have shown interesting results utilizing modifications to siRNAs. The passenger strand, when modified with a 5'-O-methyl group, prevented phosphorylation and decreased RISC activity of that strand (Chen et al., 2008). This simultaneously had the fortunate result of reducing off-target effects induced by the passenger strand. A second study found that substituting DNA nucleotides everywhere in the 5'-third of the siRNA duplex, on both the guide and passenger strands, eliminated off-target effects without substantial reduction in siRNA activity (Ui-Tei et al., 2008). Presumably, RNA at the 3'-end of the guide strand is necessary for interactions with TRBP or Ago2, either for formation of a stable RLC and RISC or for stabilizing hybridization to the target mRNA; conversely, the DNA:RNA hybrid formed at the 3'-end of the passenger strand appears not to permit the interactions that are essential for RNAi.

1.5 Approach and specific aims

The work described here aimed to develop a better understanding of the pathway of RNA interference in the effort to develop siRNAs for research and therapeutic purposes. Three aspects associated with RNAi were investigated; the theory was that if each step in the application of siRNAs could be optimized, then their overall efficacy would be maximized. The approach was to study, using experimental and computational

techniques, three key aspects of RNAi to improve the rational selection of active siRNA sequences and their subsequent delivery.

The specific aims of the present study were to:

1. Characterize the effect of secondary structure within the mRNA site targeted by the siRNA on the silencing activity of the siRNA.

Algorithms used to identify active siRNAs have historically relied on thermodynamic asymmetry and other sequence preferences for filtering, while ignoring contributions from the secondary structure of the mRNA target site. Here, certain structures predicted by UNAFold (Markham and Zuker, 2008), in which the global free energy was minimized, were found to be more amenable to silencing based on experimental results with 15 siRNAs and computational results with an additional 533 sequences. Furthermore, guide strand structure was also observed to correlate with activity, independent of mRNA structure. The methods described can readily be incorporated into improved siRNA selection algorithms.

2. Determine the characteristics of siRNAs that enhance their recognition in the RNAi pathway by the TAR RNA binding protein (TRBP).

The ability of human RISC to recognize siRNA asymmetry has been documented based on empirical evidence from experiments with hundreds of different siRNAs. However, the mechanism for the sensing remains unknown. It is shown here that TRBP

alone can perform this function. This result clarifies the role of TRBP in RNAi and suggests a possible mechanism for loading of the guide strand in RISC.

3. Develop a biocompatible and biodegradable cationic polymer nanoparticle capable of delivering siRNA into cells grown *in vitro*.

A novel polymer system generated with “click” chemistry was used to study a range of variables contributing to siRNA binding, a key step necessary for eventually achieving cellular uptake and *in vivo* silencing. High affinity binding of the nanoparticles with siRNA was found to require a combination of primary and secondary amine groups, and could be enhanced by including 12-16 carbon long alkyl side chains off the polymer backbone. Other changes in polymer functional groups also altered the affinity. Ultimately, uptake of the siRNA into cells grown in culture could be facilitated by the nanoparticles, thus providing a new platform for further development.

CHAPTER 2. IMPACT OF TARGET mRNA STRUCTURE ON siRNA SILENCING EFFICIENCY

2.1 Abstract

The selection of active siRNAs is generally based on identifying siRNAs with certain sequence and structural properties. However, the efficiency of RNA interference has also been shown to depend on the structure of the target mRNA, primarily through studies using exogenous transcripts with well-defined secondary structures in the vicinity of the target sequence. While these studies provide a means for examining the impact of target sequence and structure independently, the predicted secondary structures for these transcripts are often not reflective of structures that form in full-length, native mRNAs where interactions can occur between relatively remote segments of the mRNAs. Here, using a combination of experimental results and analyses on a large dataset, we demonstrate that the accessibility of certain local target structures on the mRNA is an important determinant in the gene silencing ability of the siRNAs. siRNAs targeting the enhanced green fluorescent protein were chosen using a minimal siRNA selection algorithm followed by classification based on the predicted minimum free energy structures of the target transcripts. Transfection in HeLa and HepG2 cells revealed that siRNAs targeting regions of the mRNA predicted to have unpaired 5'- and 3'-ends resulted in greater gene silencing than regions predicted to have other types of secondary structure. These results were confirmed by analysis of gene silencing data from previously published siRNAs, which showed that mRNA target regions unpaired at either the 5'-end or 3'-end were silenced, on average, ~10% more strongly than target regions unpaired in the center or primarily paired throughout. We found this effect to be

independent of the structure of the siRNA guide strand. Taken together, these results suggest minimal requirements for nucleation of hybridization between the siRNA guide strand and mRNA and that both mRNA and guide strand structure should be considered when choosing candidate siRNAs.

2.2 Introduction

RNA interference (RNAi) is a natural phenomenon resulting in potent and primarily specific gene silencing that occurs in most eukaryotes (Fire et al., 1998; Hannon, 2002). RNAi is initiated in cells by the presence of either short interfering RNAs (siRNAs) or microRNAs (miRNAs), which are small non-coding RNA molecules ~21 nucleotides (nt) long (Elbashir et al., 2001a). The guide strand of these small RNAs is incorporated into the active RNA induced silencing complex (RISC), which then targets mRNAs possessing regions complementary to the guide strand sequence. Upon hybridization to the target message, RISC prohibits its translation, either by cleavage of the target mRNA when guided by siRNAs, or non-degradative translational repression when guided by miRNAs. Utilization of exogenously delivered siRNAs to silence desired targets by RNAi has become a powerful tool for facilitating disease diagnosis and treatment and improving our general understanding of fundamental biological processes (Hannon, 2002; Mello and Conte, 2004).

Unfortunately, the proportion of siRNAs that are successful in repressing a target gene is low, not unlike what was found for antisense oligonucleotides (asODNs) (Stein, 1998). As with asODNs, early strategies for choosing siRNAs focused on the sequence of siRNAs, eliminating sequences based on GC content and stretches of greater than four consecutive identical bases (e.g., GGGG) (Elbashir et al., 2002). These guidelines made siRNA synthesis more convenient but were not based on mechanistic understanding. After identification of siRNAs that would be amenable to synthesis, genomic uniqueness

would be verified by BLAST searching. The remaining candidate siRNAs would then be tested for activity in cell culture to identify the best silencers.

As data from siRNA experiments has accumulated, more elaborate selection algorithms have been developed that further discriminate the most important siRNA structural and sequence features. One important design rule is based on the relative end stabilities of the siRNA duplex, with the strand whose 5'-end is more weakly hybridized incorporated preferentially into RISC (Khvorova et al., 2003; Schwarz et al., 2003; Tomari et al., 2004). Specific positional base preferences have been identified that tend to favor the differential stabilities of the two ends (Jagla et al., 2005; Reynolds et al., 2004; Ui-Tei et al., 2004). Additionally, formation of secondary structure in the siRNA guide strand can impair the ability of RISC to interact with its target mRNA (Patzel et al., 2005), analogous to what was shown with asODN (Mathews et al., 1999; Walton et al., 1999). More recently, attempts have been made to select against sequences resulting in non-specific effects, including off-target silencing and immune responses (Fedorov et al., 2006; Jackson et al., 2003; Reynolds et al., 2006; Scacheri et al., 2004; Sledz et al., 2003).

However, current siRNA selection guidelines have not typically included possible impacts of the target mRNA structure on silencing efficiency. It has been shown that siRNAs can be equally effective when targeting inside the coding region of the mRNA or the 5'- and 3'- untranslated regions (e.g., (Yoshinari et al., 2004)). Several reports have used well-defined helices to demonstrate that local target structure has a direct impact on siRNA efficacy (Ameres et al., 2007; Bohula et al., 2003; Brown et al., 2005; Far and Sczakiel, 2003; Overhoff et al., 2005; Schubert et al., 2005; Shao et al., 2007; Vickers et

al., 2003; Westerhout and Berkhout, 2007; Yoshinari et al., 2004). Gene silencing decreased when the orientation or the degree of partial base-pairing of a target construct was varied for a single siRNA (Schubert et al., 2005; Westerhout and Berkhout, 2007). This idea was extended to full-length transcripts (ICAM-1 and *survivin*) where siRNAs targeting inaccessible regions were ten-fold less active than accessible sites (Overhoff et al., 2005). Native mRNA structures likely inhibit the interaction of RISC with the target RNA, again echoing that which was found for target mRNA structure on asODN function (Vickers et al., 2000; Walton et al., 2002). One particularly intriguing indication of the importance of target mRNA structure on RNAi is illustrated by the ability of HIV-1 to overcome silencing by mutating the siRNA target sequence, doing so in a fashion that also alters its local RNA secondary structure (Westerhout et al., 2005).

In this report, we show that the influence of local mRNA target structure on the efficiency of siRNA-mediated RNAi is universal, applying to both endogenous and exogenous transcripts, across a variety of cell types. In addition, this influence can be reliably captured through prediction of the minimum free energy secondary structure of the full-length target mRNA. The impact of mRNA target structure on silencing was also found to be independent of the predicted structure of the siRNA guide strand, arguing that multiple structural factors should be taken into account when designing siRNAs.

2.3 Results

2.3.1 siRNA selection and mRNA target structure determination

We sought to establish a relationship between the gene silencing activity of siRNAs and the structure of the target mRNA. siRNAs were designed using Ambion's

siRNA Target Finder such that siRNA structures were only limited to having a 19 nt duplex with 3'-UU overhangs. No other sequence restrictions were selected. Of the 35 siRNAs returned by the algorithm, 15 were chosen for synthesis based on their GC content and the predicted structure of the target region of the mRNA (Table 2-1 and Figure 2-1). For our experiments, an EGFP reporter gene exhibiting an ~2 hr half-life was selected. This particular EGFP has been used frequently for RNAi assessment, with several effective siRNAs in existence (Kim and Rossi, 2003).

In the MFE structure, 506 out of 846 (59.8%) of the nucleotides were predicted to exist in intramolecular base-pairs. We selected our 15 designed siRNAs to sample a number of uniquely structured regions. We considered the scenario where the loop portion of the stem-loop was at the 5'-end, the center, or at the 3'-end of the siRNA target site within the mRNA (Figures 2-1 and 2-2). These regions were chosen for our classification as each has been shown to contribute uniquely to the binding and activity of RISC (Haley and Zamore, 2004), and the structures of the ends of guide strands are also known to impact silencing efficiency (Patzel et al., 2005). For central loops, we expanded the region of availability to encompass 4 nt to either side of the cleavage site because hybridization stability in this region is known to impact silencing efficiency dramatically (Elbashir et al., 2001b; Martinez and Tuschl, 2004). A loop length of 4 consecutive unpaired nucleotides was chosen based on what has been shown to be required for nucleation in bimolecular hybridization (e.g., (Hargittai et al., 2004)). For each of these three target structure types, we selected siRNAs having low (<50%), medium (50% to 60%), and high (>60%) GC content. No siRNAs with very low GC content (<30%) were returned by the algorithm, but this fact was ignored as they have

Table 2-1 Details for EGFP targeting siRNAs experimentally validated in this study.

Start	End	Antisense	Sense	GC Content (%)
71	89	ACGCUGAACUUGUGGCCGU	ACGGCCACAAGUUCAGCGU ¹	52
81	99	CUCGCCGGACACGCUGAAC	GUUCAGCGUGUCCGGCGAG	62
126	144	GAUGAACUUCAGGGUCAGC	GCUGACCCUGAAGUUCAUC ¹	48
159	177	GGGCCAGGGCACGGGCAGC	GCUGCCCGUGCCUGGCC	76
274	292	UGCGCUCCUGGACGUAGCC	GGCUACGUCCAGGAGCGCA	62
306	324	CUUGUAGUUGCCGUCGUCC	GGACGACGGCAACUACAAG	52
318	336	CUCGGCGCGGGUCUUGUAG	CUACAAGACCCGCGCCGAG	62
396	414	CAGGAUGUUGCCGUCCUCC	GGAGGACGGCAACAUCUG	57
441	459	GAUAUAGACGUUGUGGCUG	CAGCCACAACGUCUAUAUC	43
471	489	CUUGAUGCCGUUCUUCUGC	GCAGAAGAACGGCAUCAAG	48
495	513	GUUGUGGCGGAUCUUGAAG	CUUCAAGAUCGCCACAAC	48
501	519	CUCGAUGUUGUGGCGGAUC	GAUCCGCCACAACAUCGAG	52
558	576	GCCGUCGCCGAUGGGGGUG	CACCCCAUCGGCGACGGC	71
597	615	CUGGGUGCUCAGGUAGUGG	CCACUACCUGAGCACCCAG	57
639	657	CAUGUGAUCGCGCUUCUCG	CGAGAAGCGCGAUCACAUG	52

* Start sites are relative to the start codon of the EGFP coding sequence. GC Content is per 21 nt.

¹ siRNA sequences obtained from (Kim and Rossi, 2003).

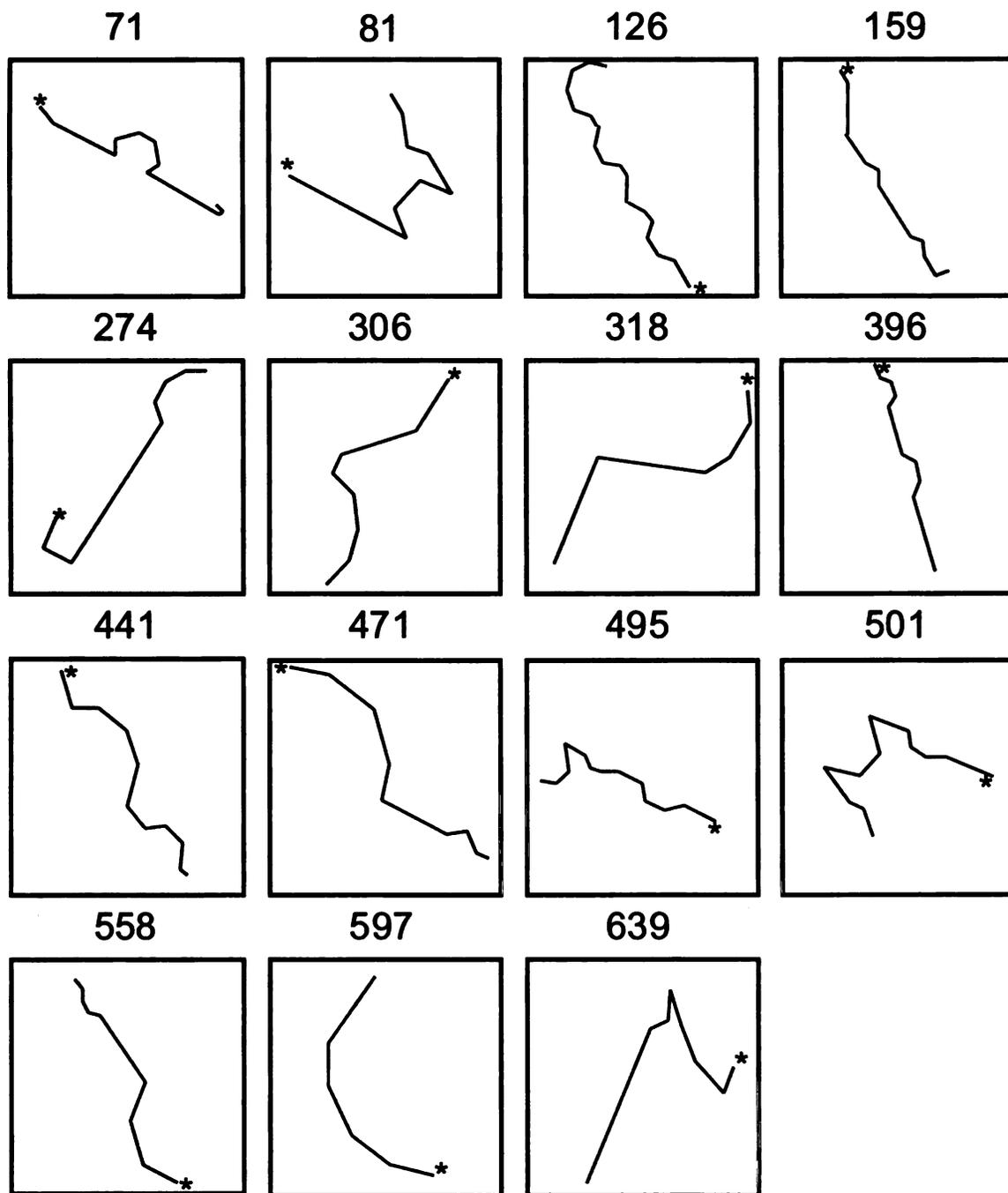


Figure 2-1 mRNA predicted structures for siRNAs targeting EGFP.

The global mRNA structures were predicted (see Appendices – Materials and Methods for Ch. 2) and the local target site (black line, * denotes 5'-end of site) is shown for each siRNA. Target start position is indicated for each of the 15 siRNAs. See (Gredell et al., 2008) for more details.

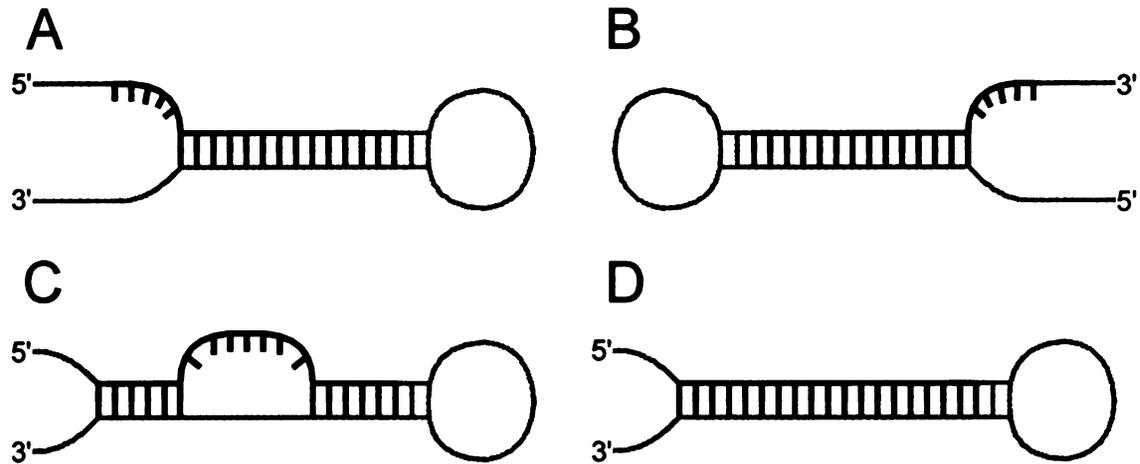


Figure 2-2 Local mRNA target structure groupings.

Four local mRNA structures were considered when grouping siRNA target regions. Classification in any group required an mRNA target region (black) that contained four consecutive unpaired nucleotides (A) at the 5'-end, '5'-loop'; (B) at the 3'-end, '3'-loop'; (C) centered around the siRNA cleavage site, 'central-loop'; or (D) nowhere along the target, 'stem'.

been shown to be less active (Reynolds et al., 2004). Search results were similarly limited for very high GC content (>70%) as only two sequences (positions 159 and 558) were identified. A lack of siRNAs with these GC contents is attributed to the overall GC content of the mRNA transcript and not necessarily to the search algorithm. To maximize the unique mRNA structures targeted, we also synthesized an siRNA predicted to target a fully looped region (pos. 597). Coincidentally, the siRNA at position 159 targeted a fully stemmed region.

2.3.2 Silencing activity of siRNAs targeting the EGFP mRNA

The activity of each siRNA was assessed in HeLa and HepG2 cells by cotransfection with the reporter EGFP plasmid. These cell lines were chosen because of their 1) popularity of use and 2) difference in transfection efficiency. No significant toxicity due to transfection was observed (data not shown). EGFP protein expression was quantified directly from live cells (Figure 2-3). Of the 15 siRNAs selected, 10 were effective in silencing the EGFP > 50% and 8 siRNAs were highly effective, resulting in silencing > 75%. These highly-effective siRNAs appeared to favor 5'- and 3'-target structures, or structures predicted to be fully looped (i.e., single-stranded). Conversely, siRNAs targeting central loops, fully hybridized regions, or regions with very high GC content showed considerably lower silencing activity on average (Figure 2-3, siRNAs 159 and 558).

2.3.3 Consideration of larger siRNA dataset – data distribution

To support our observation that the structure of the siRNA target region plays a role in gene silencing activity, we performed our target structure analysis on the siRNA

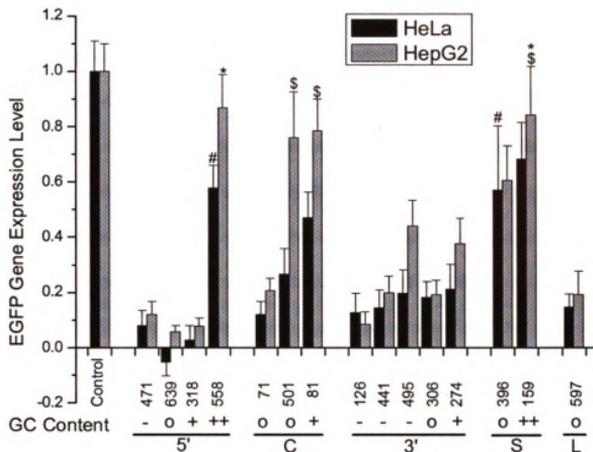


Figure 2-3 mRNA target structure dependent gene silencing.

siRNAs targeting the EGFP mRNA (see Table 2-1 and Figure 2-1) were cotransfected into HeLa and HepG2 cells and EGFP fluorescence levels were measured after 24 hrs. siRNAs are grouped according to the predicted structure (Figure 2-2; 5' – 5'-loop; C – central loop; 3' – 3'-loop; S – stem; L – completely unpaired, 'loop') of their target sites and by GC content ('-' – <50%; 'o' – 50-60%; '+' – 60-70%; '++' – >70%). Values represent mean ± standard deviation for at least eight independent experiments (n≥8). All siRNA treatments were significantly different from Control (Student's t-test; p<0.01) except those denoted by * for HepG2 cells. Comparisons were also made for each siRNA relative to the sequence giving the smallest change in EGFP expression level for the corresponding cell line. HeLa: all siRNA treatments are significantly different from siRNA 159 except those denoted by # (p<0.01). HepG2: all siRNA treatments are significantly different from siRNA 558 except those denoted by \$ (p<0.01).

silencing database from (Shabalina et al., 2006). We analyzed the distribution of data within this dataset and found that the distribution of siRNAs was heavily skewed to siRNAs with high activity (low gene expression level) (Figure 2-4). Further investigation revealed that this was partially a result of incorporating siRNAs from (Reynolds et al., 2004), where 180 siRNAs targeting every other position of the cyclophilin B and firefly luciferase were used. Removal of these siRNAs from the dataset rendered the data relatively uniformly distributed, as would be expected for a well-sampled dataset. Our subsequent analyses were performed with and without the Reynolds data so as to characterize any uniqueness associated with this particular set of siRNAs.

2.3.4 Distribution of base-pairs within the mRNA target and the effect on silencing activity

MFE mRNA secondary structures were predicted for the complete sequence of each gene, as published by NCBI, in the dataset using UNAFold v. 3.4, with default settings (Markham and Zuker, 2005; Zuker, 2003). These structures were used to determine the influence of the total number of base-pairs in the mRNA region targeted by siRNA. We observed that siRNAs targeted structures ranging from those completely unpaired (zero base-pairs in the mRNA target region) to those fully paired (19 base-pairs in the target; Figure 2-5A), with the majority of target regions containing between 10 and 16 base-pairs. The average silencing efficiency tended to decrease as the total number of base-pairs increased up to 16, beyond which the number of siRNAs in each group is too limited to assess any significant trend (Figure 2-5B).

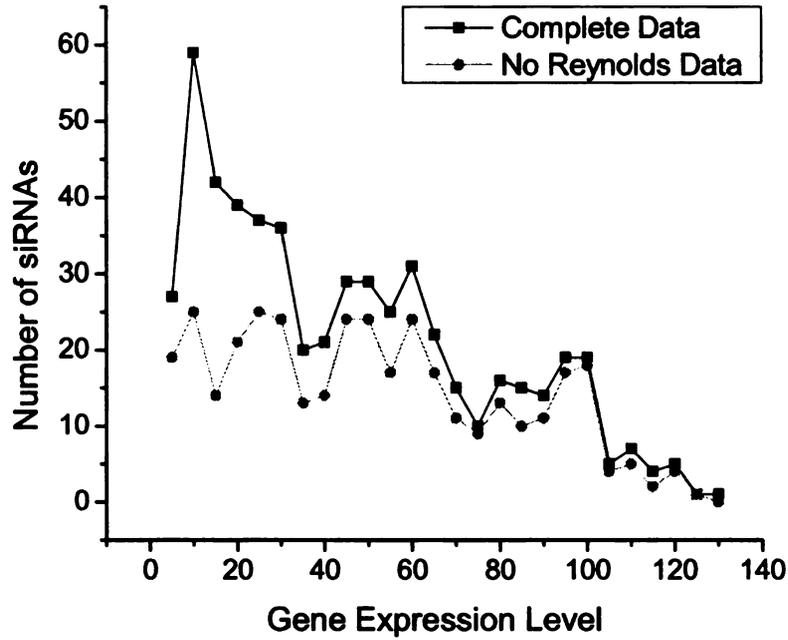


Figure 2-4 Distribution of siRNAs.

The frequency of siRNAs giving an average gene expression remaining after siRNA treatment (grouped into bins of 5%; e.g., 0-5, 5-10, etc.) is shown for the siRNAs from the complete dataset and after the data from (Reynolds et al., 2004) was removed.

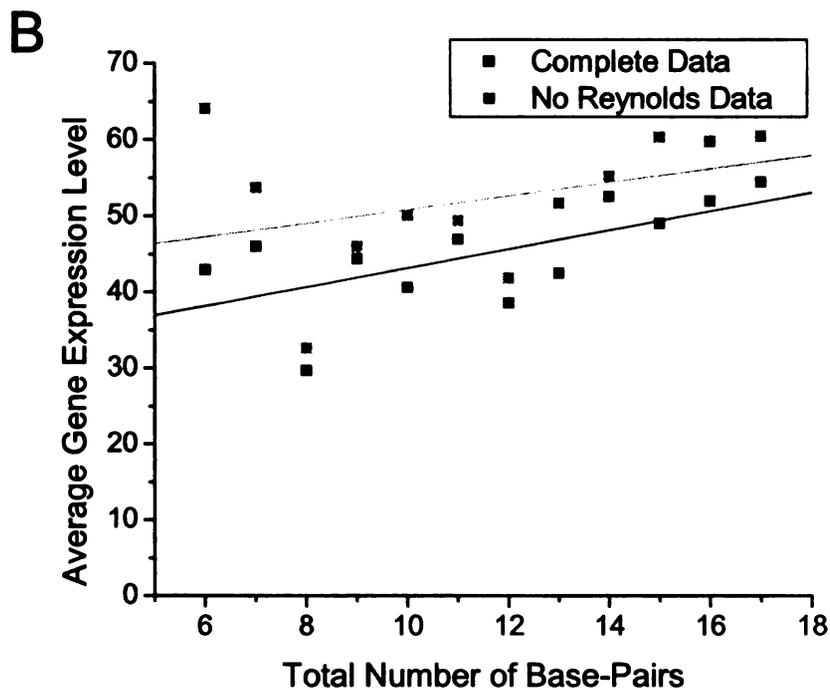
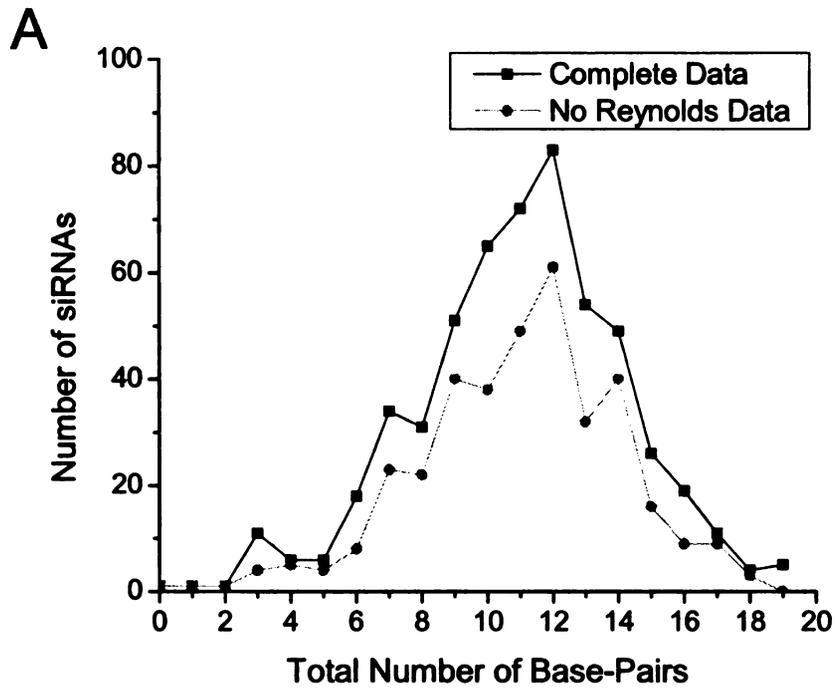


Figure 2-5 Influence of the number of base-pairs within the target site.

Shown are (A) the frequency of siRNAs targeting mRNA sites with the predicted number of base-pairings and (B) the effect of the number of base-pairs in the target site on the average gene expression level remaining after siRNA treatment. For (B), only points where more than 15 siRNAs were tested are shown. The linear regressions show a positive correlation. As in Figure 2-4, data are plotted with and without the data from (Reynolds et al., 2004).

2.3.5 mRNA target loop size and location effects

Since the total number of base-pairs within an mRNA target appeared to influence the degree of silencing, we hypothesized that the number and location of unpaired nucleotides would be equally influential in gene silencing. Using the large dataset, we determined the location of unpaired loops in the mRNA target sites. We considered scenarios where loops contained 1, 2, 3, 4, or 5 consecutive unpaired nucleotides and called this the “window size”. The window was walked along the target site (5'-end of mRNA target is position 1), and the average silencing activity for those siRNAs that were unpaired in that window was calculated (Figure 2-6). Even at a window size of 1, the silencing activity tended to be > 5% better at the 5'- and 3'-ends than at the center, supporting the results we obtained in our experimental system (Figure 2-3). This trend became more pronounced as the window size increased from 1 to 4 where the difference in silencing activity from ends to center was 8-10%. It is worthwhile to note that 4 available nucleotides has been shown to permit nucleation of nucleic acid hybridization (Hargittai et al., 2004), supporting our result. Little change was observed increasing from a window size of 4 to 5 (Figure 2-6). We therefore defined our loop size for structure classification as four consecutive unpaired nucleotides (W=4) and refined our target structure matrix (Table 2-2).

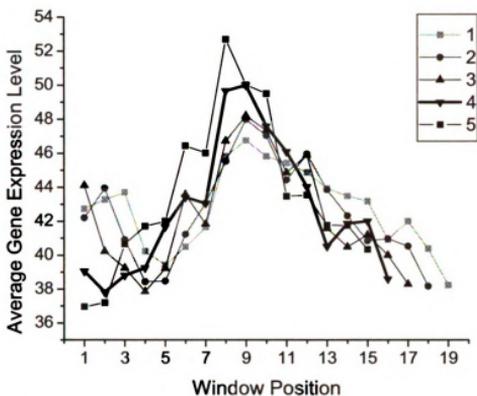


Figure 2-6 Effect of window size on gene expression level.

The profiles of the average gene expression level are shown for siRNAs (complete dataset) that were completely unpaired in the reading window ($W=1, 2, 3, 4,$ or 5 consecutive nucleotides). Because the window size changes, the number of possible windows changes concomitantly; that is, there are 19 windows of size 1 but only 16 windows of size 4 along a 19 nt long siRNA.

Table 2-2 Definition of RNA structures and siRNA activity within each group.

Structure Type	Location of Unpaired Nucleotides	Structured mRNA				Structured Guide Strand			
		# of siRNAs		Average Gene Expression Level		# of siRNAs		Average Gene Expression Level	
		Complete Data	No Reynolds	Complete Data	No Reynolds	Complete Data	No Reynolds	Complete Data	No Reynolds
5'	1-5	121	84	39.4	44.8	233	159	39.4	44.7
Central	7-13	98	60	47.2	51.7	119	70	48.1	52.8
3'	15-19	123	85	40.3	42.9	241	170	44.0	49.6
Stem	None	144	98	48.4	55.5	9	5	52.5	68.5
Other	3-9 or 11-17	88	53	41.6	50.8	N/A	N/A	N/A	N/A

Target structures were defined when 4 or more consecutive unpaired nucleotides were present in the specified regions. Nucleotide positions are relative to the 5'-end of the target mRNA. See Figure 2-2 for illustration. N/A, not applicable.

For targets that contained both a central loop and either a 5'- or 3'-loop, we classified these as one of the latter. The distribution of the data was similar in each group, with noticeably less skew when ignoring the Reynolds data (data not shown). The 5'- and 3'-loop structures showed improved silencing activity, while siRNAs targeting central loops demonstrated reduced functionality. Furthermore, stem structures gave ~12% lower silencing efficiency as compared to 3'-loops. The functionality of siRNAs within these groups (Table 2-2) was then compared to each remaining group using an unpaired Student's t-test. The silencing activity of siRNAs in either the 5'- or 3'-groups was significantly higher than for siRNAs in the center, stem, and other groups (Table 2-3); the p-value for the combined 5'- and 3'-groups versus all others was 0.0038 and 0.0005 for the Complete Data and No Reynolds Data, respectively, providing statistical support to our group classifications.

2.3.6 siRNA guide strand structure effects

To demonstrate that siRNA guide strand structures were not confounding our analyses of mRNA local target structures, we determined the MFE secondary structure for each guide strand and performed the same analysis as described above for the mRNA. Guide structures of the 5'- and 3'-loop types (defined using $W=4$) were twice as common as central loops, with <2% forming stems; no guide strands formed “other” structures (Table 2-2). Consistent with a previous report (Patzel et al., 2005), guide structures with 5'-loops resulted in improved gene silencing relative to central and 3'-loops (Table 2-2), as confirmed by Student's t-test (Table 2-4).

Table 2-3 Statistical analysis of siRNA activity between groups for mRNA target structures.

Complete Data					
	Central	3'	Stem	Other	All
5'	0.0281	0.4101	0.0100	0.3007	0.0503
Central		0.0472	0.3812	0.1095	0.0853
3'			0.0192	0.3772	0.1039
Stem				0.0589	0.0150
Other					0.2834
5' and 3'					0.0038

No Reynolds Data					
	Central	3'	Stem	Other	All
5'	0.0867	0.3439	0.0094	0.1321	0.0822
Central		0.0446	0.2267	0.4410	0.2075
3'			0.0034	0.0752	0.0232
Stem				0.1940	0.0068
Other					0.3028
5' and 3'					0.0005

The silencing activities for siRNAs in the group in the left-hand column were compared to the activities for siRNAs in the group along the top row and the resulting p-value is shown. Note that the p-value for the 5' versus 3' comparison did not take into account the fact that ~5% of the siRNAs can be classified as both 5'- and 3'-loops; therefore the data are not completely independent and thus those p-values are larger than if the groupings were independent. However, in the comparison of siRNA activities for groups in the left-hand column versus activities from all other siRNAs not in that group ("All" group in the top row), only independent siRNAs were considered. We also compared siRNAs in either the 5' or 3' groups together versus "All". This p-value was lower than either the 5' versus "All" or the 3' versus "All" p-values because the 5' and 3' groups were not significantly different from each other, yet they were included in the "All" grouping.

Table 2-4 Statistical analysis of siRNA activity between groups for guide strand structures.

Complete Data

	Central	3'	Stem	All
5'	0.0058	0.0491	0.1566	0.0029
Central		0.1255	0.3645	0.0363
3'			0.2539	0.3730
Stem				0.2372
5' and 3'				0.0227

No Reynolds Data

	Central	3'	Stem	All
5'	0.0331	0.0718	0.1264	0.0129
Central		0.2347	0.2168	0.1123
3'			0.1749	0.3105
Stem				0.1619
5' and 3'				0.0583

The silencing activities for siRNAs in the group in the left-hand column were compared to the activities for siRNAs in the group along the top row and the resulting p-value is shown. Note that the p-value for the 5' versus 3' comparison did not take into account the fact that ~30% of the siRNAs in the guide strand structure analysis can be classified as both 5'- and 3'-loops; therefore the data are not completely independent and thus those p-values are larger than if the groupings were independent. However, in the comparison of siRNA activities for groups in the left-hand column versus activities from all other siRNAs not in that group ("All" group in the top row), only independent siRNAs were considered. We also compared siRNAs in either the 5' or 3' groups together versus "All". However, the 5' group, and not the 3' group, was significantly different than the "All" group. Therefore the significance in the comparison of 5' and 3' versus "All" is due to the inclusion of the 5' group results and not the 3' group.

We next considered a pair-wise comparison of each mRNA structure group versus each guide structure group (Table 2-5). siRNAs with any combination of 5'- or 3'-loops in the mRNA or guide structures tended to give the best silencing, with >5% difference in activity compared to combinations including central, stem, or other groups. siRNAs with central loops in the guide strand or targeting stems in the mRNA had lower silencing activities, as was expected. Taken together, these results suggest that inclusion of basic mRNA secondary structural information, in parallel with siRNA guide strand structure details, can improve the likelihood of identifying active siRNAs.

2.4 Discussion

Accurately accounting for significant controlling parameters in siRNA design continues to pose a considerable problem for RNAi applications. Ideally, all mRNAs targeted for cleavage by RISC would be entirely single-stranded and free from ribosomes or other bound molecules, allowing for uninhibited hybridization of any siRNA guide strand to its complementary target. Even ignoring the presence of ribosomes, it has been shown that the accessibility of the target mRNA through complex secondary structures can influence RNAi (Ameres et al., 2007; Bohula et al., 2003; Brown et al., 2005; Long et al., 2007; Overhoff et al., 2005; Schubert et al., 2005; Shao et al., 2007; Vickers et al., 2003; Westerhout and Berkhout, 2007; Yoshinari et al., 2004). In this work, our goal was to facilitate incorporation of target mRNA structure into siRNA design algorithms by investigating which mRNA structure types were most amenable to silencing. We experimentally showed that specific regions of the target mRNA were more susceptible to RNAi silencing than others. Furthermore, these results were in agreement with computational analyses of data available in the literature. The majority of our results (41

Table 2-5 Distribution of siRNA functionality for each target structure classification.

		Guide Strand			
		5'	Central	3'	Stem
mRNA	5'	35.4 0.4996 43	43.2 0.0141 24	40.9 0.0876 64	18.9 4
	Central	41.6 0.0149 31	49.9 0.4147 41	50.7 0.1948 29	56.0 1
	3'	35.9 0.3969 64	49.5 0.0258 19	36.7 0.1405 50	70.2 4
	Stem	45.0 0.0035 66	52.1 0.4659 23	49.9 0.1011 71	N/A 0

Target structures were defined when 4 or more consecutive unpaired nucleotides were present in the specified regions of either the siRNA guide strand or the mRNA target. The entries in each box represent the average gene expression level (top), the p-value from a one-sided Student's t-test (middle), and the total number of siRNAs within that group (bottom) for siRNAs in the Complete Dataset.

of 42 genes) are based on the silencing of endogenous mRNAs, thus avoiding any artifacts that may have arisen by silencing only exogenous or engineered constructs. Despite the variety of systems and readouts used to obtain it, the data clearly showed that siRNAs targeting regions unpaired at either the 5'- or 3'-end silenced, on average, 8% better than siRNAs targeting regions unpaired in the center or without any unpaired windows. Though this net effect could be argued to be small, our observation is consistent with a recent report showing an average difference in knockdown of ~14% for 101 shRNAs when accounting for an energy parameter in target accessibility calculations (Shao et al., 2007). It is noteworthy that in this same report a difference of ~35% was found when only considering shRNAs that have a favorable duplex asymmetry. While we did not take into account duplex asymmetry in our analyses, these results strongly suggest that our differences in average silencing would be improved by removing siRNAs that are not asymmetric. Regardless, provided that it comes at low computational and time costs, the additional information gained by using mRNA target structure predictions is warranted for incorporation into siRNA design algorithms to enhance the likelihood of selection of active siRNAs.

As the number of experimentally tested siRNAs increases, it becomes vitally important to note the conditions for both design and application of effective *and ineffective* sequences when reporting results. By doing so, one could greatly enhance the set of siRNAs available for bioinformatics studies and development of design algorithms (Matveeva et al., 2007). Algorithms would be similarly assisted by using silencing results from siRNAs randomly selected for their target, since more elaborate selection methods may bias sequences towards parameters already shown to be relevant, such as

GC content and other particular positional base preferences (Reynolds et al., 2004). We therefore designed our siRNAs targeting EGFP using an algorithm developed to take advantage of a simple method for enzymatic siRNA synthesis (http://www.ambion.com/techlib/misc/siRNA_finder.html). This algorithm only required that the target site lie immediately downstream of 'AA' dinucleotides, and so all of our experimental sequences contained UU overhangs (making the entire 21 nt of the siRNA complementary to the target mRNA). This is certainly a potential source of bias, and may be one reason why our initial experiments yielded such a high proportion of active siRNAs (8 out of 15 silenced the EGFP more than 75%). It is noteworthy that the siRNA targeting a fully unpaired region resulted in >75% silencing (Figure 2-3; target 597) and that the siRNAs targeting fully paired regions (targets 159 and 396) yielded only ~20-40% reduction, an expected trend for the most and least ideal target types, respectively. Moreover, the total number of base-pairs within the mRNA target site, another metric for target site stability, showed a normal distribution around 11 base-pairs, with or without the data from (Reynolds et al., 2004) (Figure 2-5A). This suggests that most siRNA target sites have at least half of the target region sequestered in native structure. As with the loop and stem targeting siRNAs, siRNAs targeting regions with fewer total base-pairs tended to reduce gene expression more effectively (Figure 2-5B).

When analyzing data for parametric information, it is important that the dataset is as unbiased as possible and contains enough data points to draw the appropriate conclusions. Beginning with the entire database of (Shabalina et al., 2006), we pared it down to only those siRNAs targeting human genes shorter than 6,000 nts. This dramatically increased our structural prediction speed while only reducing our dataset by

18%, from 653 to 533 siRNAs. After examining the activity profile of the remaining sequences, we observed that the data were skewed to siRNAs that reduced gene expression levels below 35% (Figure 2-4), due to the Reynolds data. Despite our concerns about the unequal weighting in the data, inclusion of these data in our analyses did not significantly alter the results for 5'-, central, or 3'-loops (Figures 2-5 and 2-6, Tables 2-2 and 2-3). Some effect, though, was seen on the sequences in the "other" grouping. "Other" refers to those sequences that have loops but not in regions to be classified in the 5'-loop, 3'-loop, or central class. The average gene expression remaining for this group increased from 42% to 51% upon removal of the Reynolds data. Similarly, the analyses implemented in Table 2-3 (and data not shown) revealed no statistically significant effect from this group. These results indicate that structures that fall into the "other" class are not as critical for defining silencing efficiency, though this issue is still up for debate (Katoh and Suzuki, 2007; Schwarz et al., 2006).

Predicted secondary structures for RNA have been determined using algorithms such as mfold (Mathews et al., 1999; Zuker, 2003) or the Vienna RNA Package (Hofacker, 2003) that utilize nearest-neighbor energies to obtain an MFE structure and a set of suboptimal structures. However, others have suggested that this unnecessarily assumes that the MFE structure is the most prevalent structure of the mRNA in the cell (Ding et al., 2004). Instead, an ensemble of foldings encompassing a statistically significant sampling (~1000 structures) is considered more appropriate (Ding et al., 2004). This lends itself to a stochastic approach where the probabilities of interactions can be assessed. Both methods have proven useful in attempting to account for mRNA target structure in RNAi (Ameres et al., 2007; Heale et al., 2005; Long et al., 2007;

Overhoff et al., 2005; Patzel et al., 2005; Schubert et al., 2005; Shao et al., 2007; Westerhout and Berkhout, 2007). While we did not consider ensembles in this work, for design purposes, our results (Figure 2-3 and Tables 2-2 and 2-5) and those of several others (Far and Sczakiel, 2003; Luo and Chang, 2004; Patzel et al., 2005; Schubert et al., 2005; Westerhout and Berkhout, 2007) show that mRNA MFE structures determined by mfold alone provide information that is useful in refining the pool of candidates for selection of active siRNAs.

It is well-established that RNA secondary structure predictions worsen as the length of the sequence to be folded increases. It is therefore common to predict folded structures for sequences of a given length, e.g., < 700 nt (Doshi et al., 2004; Mathews et al., 2004). For predictions of local structures, the folded sequence is generally specified to encompass a given length (~100 nt) of sequence upstream and downstream of the target (Heale et al., 2005; Long et al., 2007). This approach, though possibly valuable for gleaning additional information about the predicted target region structures, would not be of interest to individuals designing siRNAs due to the computational expense, the number of folds to analyze, and the lack of a rigorous definition as to the appropriate sequence length to fold. As such, we did not investigate the use of a segmented approach for prediction of our target structures.

The extent to which certain regions of the siRNA guide strands, and consequently the mRNA regions targeted by those strands, influence RNAi has been somewhat controversial (Haley and Zamore, 2004; Long et al., 2007; Patzel et al., 2005; Shao et al., 2007; Westerhout and Berkhout, 2007). One report found that the 5'-end of the siRNA guide strand contributes more to the binding of the target mRNA than the center and 3'-

end, which affect the helical geometry of the siRNA/mRNA hybrid (Haley and Zamore, 2004). This conclusion is consistent with our analysis showing a preference for unpaired 5'-ends in the guide strand over both central and 3'-loops (Tables 2-4 and 2-5). Our guide strand results support those of Patzel and colleagues who found that for guide strands forming stem-loop structures the 5'-end was more influential than the 3'-end, with either being more critical than an unpaired central region (Koberle et al., 2006; Patzel et al., 2005). These results imply that 3'- and 5'-ends of the mRNA regions targeted by the guide strands are also important. This is further supported by a recent report that describes accessibility of the 3'-end of the target site as an important determinant of RISC activity (Ameres et al., 2007). It has also been shown using a reporter construct that a free 3'-end in the target region improved RNAi silencing (Westerhout and Berkhout, 2007), though only limited structures were considered which perhaps masked the contributions from the 5'-end. Destabilizing native mRNA structure both upstream and downstream of the siRNA target sequence can enhance the association rate of RISC as well (Ameres et al., 2007; Brown et al., 2005), which would suggest an improved likelihood of silencing. Another recent study found that improved silencing occurs at target sites bordered on both the 5'- and 3'-ends by regions of high AU content (Nielsen et al., 2007), as these would presumably have relatively weaker native mRNA structure. However, two other reports reached no such conclusions, instead noting that siRNA/mRNA nucleation can occur anywhere along the target site (Long et al., 2007). These papers used a threshold energy term in the analysis of miRNA (Long et al., 2007; Shao et al., 2007) and siRNA (Shao et al., 2007) binding to structured targets. After nucleation, a second energy parameter regarding helix elongation was required to fully

describe the miRNA behavior. This report (Long et al., 2007) also showed that nucleation of four consecutive unpaired nucleotides gave better correlations to gene inhibition. This latter result is in agreement with our observations of siRNA initiated gene silencing (Figure 2-6). Unfortunately, our attempts to describe the energetics of the siRNA/mRNA interaction using the RNAup algorithm from the Vienna RNA Package (Muckstein et al., 2006) or the algorithm of (Heale et al., 2005) were unsuccessful (Figure 2-7). The fact that an effect was observed at both the 5'- and 3'-ends of the target suggest that the seed site (i.e., the 5'-end) and the relative differential stability between the 5'- and 3'-ends are not the only important parameters for si/miRNA functionality. Most likely, it is easier for mRNA targets with four consecutive unpaired nucleotides at one or both ends to form a stable hybrid with a guide strand also lacking structure at one or both ends. Once initiated, the dsRNA helix can elongate, possibly facilitated by the helicase component associated with active RISC (Robb and Rana, 2007).

Using a broad analysis of siRNA-mediated RNAi from the literature, we have shown that siRNAs targeting mRNAs with predicted regions of four consecutive unpaired nucleotides at either the 5'- or 3'-ends of the target site are more potent for inducing RNAi-based gene silencing than when the center is unpaired or when the target site has no unpaired regions. Additionally, these observations are consistent with the mechanistic understanding of guide strand loading in the RNAi pathway. As these results were based on predicted MFE structures, they demonstrate the utility of mRNA secondary structure predictions in enhancing the likelihood of identifying active siRNAs.

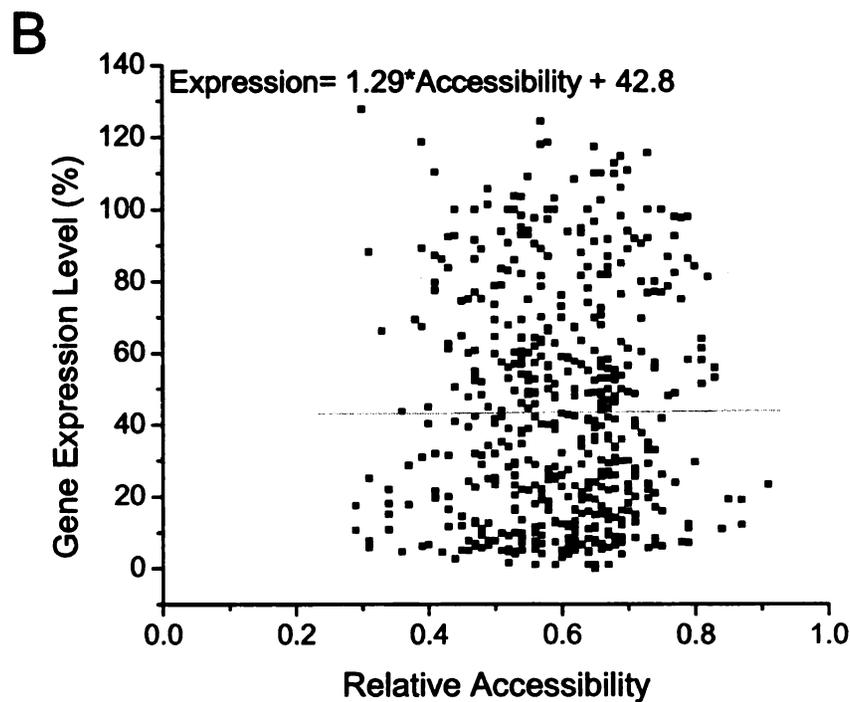
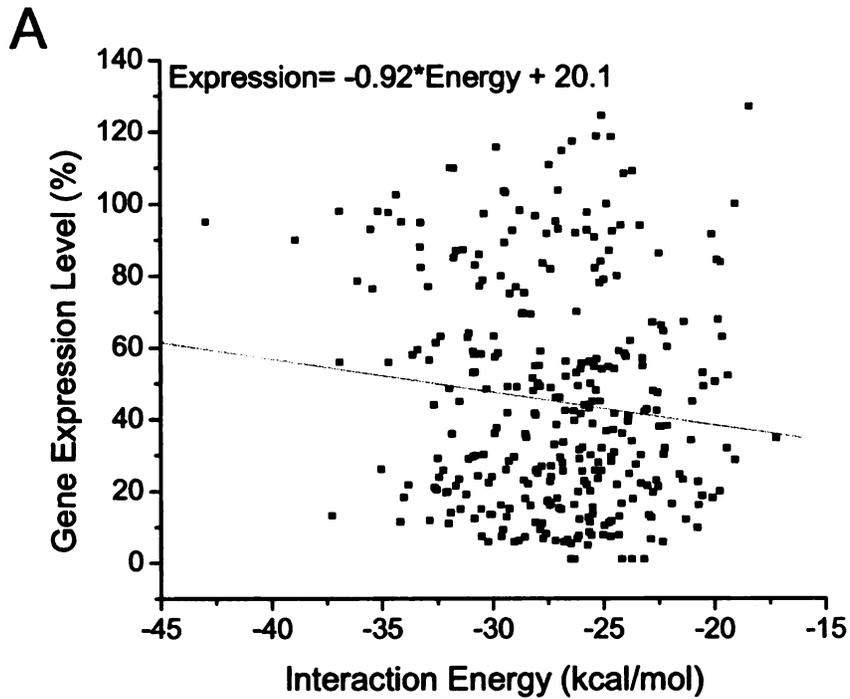


Figure 2-7 siRNA-mRNA interaction predictions.

(A) The interaction energies were calculated by the RNAup algorithm from the Vienna RNA Package. The linear regression (grey line) distribution coefficient (R) is 0.1190.

(B) The relative accessibilities of the mRNA were calculated by the algorithm of Heale, et al. (2005). The linear regression (grey line) distribution coefficient (R) is 0.0045.

CHAPTER 3 RECOGNITION OF siRNA BY TAR RNA BINDING PROTEIN (TRBP)

3.1 Abstract

The recognition of small interfering RNAs (siRNAs) by the RNA induced silencing complex (RISC) and its precursor, the RISC loading complex (RLC), is a key step in the RNA interference pathway that controls the subsequent sequence-specific mRNA degradation. In *Drosophila*, selection of the appropriate guide strand has been shown to be mediated by the RLC protein R2D2, which senses the relative hybridization stability between the two ends of the siRNA. A protein with similar function has yet to be conclusively identified in humans. We show here that human TAR RNA binding protein (TRBP) alone can bind siRNAs *in vitro* and sense their asymmetry in the absence of ATP. We also show that TRBP can bind 21 nt ssRNA, though with far lower affinity than for double-stranded siRNA, and that the binding reflects the bias observed with the full siRNA duplex. This suggests that TRBP binding may be both sequence and stability dependent. A computational analysis of published silencing results supports this hypothesis. Together these results demonstrate the importance of the siRNA-TRBP interaction in the formation of active RISC and in siRNA guide strand selection in RNAi.

3.2 Introduction

RNA interference (RNAi) is a means of enacting specific silencing of the expression of a target gene (Fire et al., 1998; Hannon, 2002). The pathway is initiated when long dsRNAs or pre-miRNAs are processed by the RNase III family enzyme Dicer into 21-27 nt long siRNAs or miRNAs with 5'-phosphates and 3'-dinucleotide overhangs (Bernstein et al., 2001). The resulting small noncoding RNA is then loaded into the RNA induced silencing complex (RISC) (Maniataki and Mourelatos, 2005). Typically, RISC cleaves the target mRNA when using an siRNA to guide sequence complementarity, whereas miRNAs tend to result in translational inhibition (Siomi and Siomi, 2009). RISC is minimally composed of Dicer, the TAR RNA binding protein (TRBP), and Argonaute 2 (Ago2), although several other proteins have been shown to associate with RISC *in vivo* (Chendrimada et al., 2005; Gregory et al., 2005; Haase et al., 2005; Kok et al., 2007; Lee et al., 2006; MacRae et al., 2008; Robb and Rana, 2007). Dicer appears to couple dsRNA processing with selection of one strand of the siRNA for loading onto Ago2 (Kim et al., 2005; Rose et al., 2005; Siolas et al., 2005), the catalytic component of RISC responsible for degradation of the opposite siRNA strand and the target mRNA (Liu et al., 2004; Matranga et al., 2005; Meister et al., 2004; Rand et al., 2005; Rivas et al., 2005; Song et al., 2004). The precise contribution of TRBP to this process has yet to be established.

TRBP was first identified by its ability to bind the TAR RNA structure present in human immunodeficiency virus (HIV-1) transcripts (Gatignol et al., 1991). Later, it was shown that TRBP can inhibit Protein Kinase R (PKR), an important contributor to the

innate response to viral infection (Benkirane et al., 1997). Most recently, it was found that TRBP interacts directly with Dicer, as well as the activator of PKR (PACT), through the Medipal domain present at the C-terminus of TRBP (Laraki et al., 2008), and that the interaction is not mediated solely by dsRNA (Haase et al., 2005). A similar interaction is detected in *Drosophila*, where Dicer-2 associates with the dsRNA-binding protein (dsRBP) R2D2 (Liu et al., 2003). Together, the Dicer-2/R2D2 complex senses the relative thermodynamic asymmetry within an siRNA (Tomari et al., 2004). R2D2 binds to the more stable (dsRNA-like) end, leaving Dicer-2 to bind the opposite end. Furthermore, binding is enhanced by the presence of a 5'-phosphate on the passenger strand; hydroxyl groups, such as those present during chemical siRNA synthesis, inhibit binding (Tomari et al., 2004). As a result of the biased binding, a specific strand of the siRNA, termed the guide strand, is preferentially loaded into RISC, and the complementary passenger strand is degraded (Matranga et al., 2005; Rand et al., 2005).

Human RISC has been extensively shown to sense asymmetry between the ends of the siRNA or miRNA, leading to the preferential incorporation of a single strand as the guide strand (Gregory et al., 2005; Khvorova et al., 2003; Maniataki and Mourelatos, 2005; Reynolds et al., 2004). This has been accomplished with crude cell extracts and immunopurified proteins (Gregory et al., 2005; Maniataki and Mourelatos, 2005), and more recently using only recombinant Dicer, TRBP, and Ago2 in a 1:1:1 stoichiometry (MacRae et al., 2008). The effect has also been observed empirically during the analysis of large siRNA or miRNA datasets (Khvorova et al., 2003; Reynolds et al., 2004), and is currently the cornerstone for predicting siRNA efficacy (Lu and Mathews, 2007; Shao et al., 2007). However, unlike in *Drosophila*, no studies using human cells or proteins have

conclusively shown how the asymmetry is detected. Since Ago2 alone cannot bind and select the siRNA guide strand (Rivas et al., 2005), it suggests that its associated partners Dicer and/or TRBP are necessary and capable of initial siRNA loading, even though other factors such as RNA helicase A (RHA) or PACT may contribute to the effect *in vivo* (Kok et al., 2007; Lee et al., 2006; Robb and Rana, 2007).

We were interested in characterizing the binding of siRNAs by recombinant TRBP *in vitro* to determine whether and how it would sense asymmetry in siRNAs. Our results show that TRBP binds siRNAs and not the analogous siDNAs or siDNA/RNA hybrids. Furthermore, we demonstrate that TRBP alone can sense the asymmetry of an siRNA duplex, thus establishing a key function of RNAi in humans. Our data also show that TRBP-siRNA binding may be somewhat terminal sequence dependent, a result which was supported through analysis of a large-scale silencing dataset. Additionally, ssRNAs were recognized with a bias that reflected the asymmetry detected in the full duplex, providing further support for some sequence dependence of TRBP-RNA binding.

3.3 Results

To study the contribution of TRBP to siRNA asymmetry sensing in humans, we expressed and purified recombinant human TRBP as a fusion product with maltose binding protein (MBP), as well as MBP alone for control (Figure 3-1). Both plasmids were generously provided by Professor Anne Gatignol (Daviet et al., 2000; Laraki et al., 2008). Purified MBP and TRBP products were resolved by denaturing gel electrophoresis and visualized by protein staining or western blotting (Figure 3-1A-C). Proper function of the TRBP was then characterized by its ability to bind TAR RNA using a native gel shift assay (Figure 3-1D). Multiple shifted complexes were visible,

consistent with previously observed multimerization of TRBP (Cosentino et al., 1995; Daviet et al., 2000).

3.3.1 TRBP recognition of small nucleic acids

TRBP has previously been shown to bind siRNAs (Kato and Suzuki, 2007; Parker et al., 2008; Ui-Tei et al., 2008). We also found that to be the case in a native gel shift assay (Figure 3-2, lanes 1 and 2). TRBP formed at least four distinct shifted complexes with the siRNA, corresponding approximately to complexes containing at least one siRNA molecule in a complex with one, two, four, or ~six TRBP molecules (Figure 3-3). As expected, larger complexes became more prominent at higher TRBP concentrations. Furthermore, although TRBP can bind long ssRNAs (Gatignol et al., 1993), no binding was observed for the ssRNA corresponding to the antisense strand (AS) of the duplex siRNA at the TRBP concentration tested (Figure 3-2, lanes 3 and 4). In addition, no binding was found for an siDNA or the corresponding single-stranded AS siDNA counterpart (Figure 3-2, lanes 5-8). We next tested DNA-RNA hybrids as these have been shown to be RNAi-competent (Hohjoh, 2002; Lamberton and Christian, 2003; Ui-Tei et al., 2008). Somewhat surprisingly, no binding was detected with the heteroduplexes, either. Repeating the experiment with MBP alone confirmed that any observed binding was due to TRBP and not non-specific binding by the MBP portion of the fusion product (Figure 3-4). These results are consistent with dsRBPs, such as TRBP, PKR, and PACT, requiring A-form, double-stranded regions for binding.

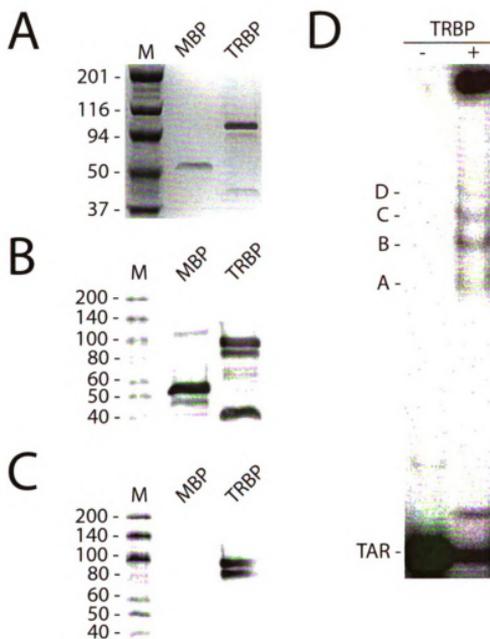


Figure 3-1 Characterization of recombinant TAR RNA binding protein (TRBP).

Recombinant TRBP was prepared as a fusion product with maltose binding protein (MBP) as described (Materials and Methods). (A) Purified MBP or TRBP proteins were resolved by SDS/PAGE and visualized by staining with Gel Code Blue. (B and C) Western blots of gels as shown in (A) were performed with antibodies specific for (B) MBP or (C) TRBP. (D) Gel mobility-shift analysis of TRBP with ³²P-labelled TAR RNA was achieved by incubating 1000 nM of protein with a limiting amount of TAR RNA (10⁴ cpm; <1.0 nM) and resolving protein-RNA complexes by native PAGE. Four distinct complexes were visible (bands A-D). M = protein size marker (kDa).

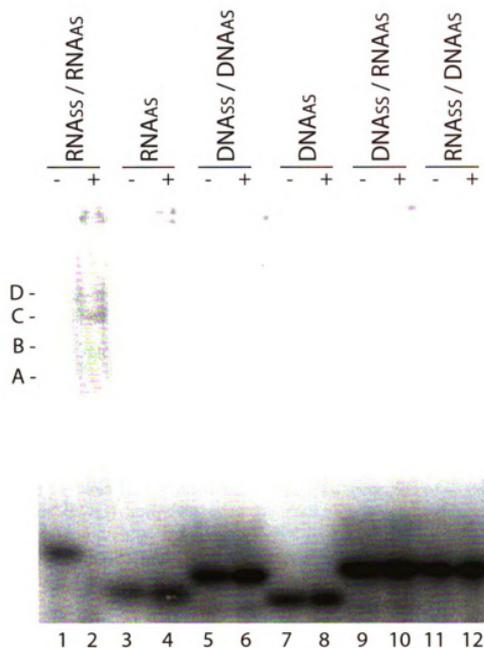


Figure 3-2 TRBP binding preferences for nucleic acids.

Native gel shift assay of 32 P-labeled siRNA (lanes 1&2), ssRNA (3&4), siDNA (5&6), ssDNA (7&8), and DNA/RNA (9&10) or RNA/DNA (11&12) hybrids with (+) or without (-) 2000 nM TRBP. Shifted TRBP-siRNA complexes are indicated by A (monomer of TRBP with siRNA), B (dimer of TRBP), C (tetramer of TRBP), and D (larger order structure of TRBP).

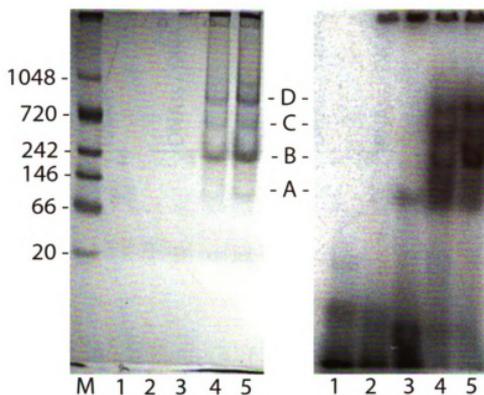


Figure 3-3 TRBP multimerization on siRNA.

Native gel shift assay of a ^{32}P -labeled siRNA with increasing concentrations of TRBP (0 to 2000 nM). The gel was first stained with Gel Code Blue for total protein (left; M = protein size marker (kDa)) and then used for radio-imaging (right). Shifted TRBP-siRNA complexes are indicated by A (monomer of TRBP with siRNA), B (dimer of TRBP), C (tetramer of TRBP), and D (larger order structure of TRBP).

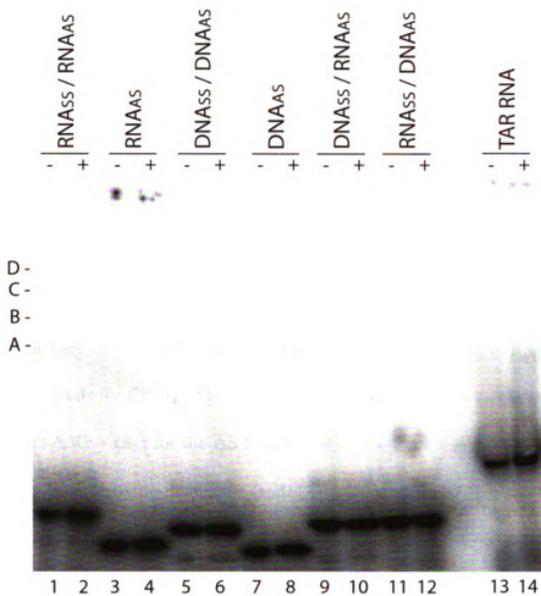


Figure 3-4 No MBP binding of small nucleic acids.

Native gel shift assay as described in Figure 3-2 but with TRBP replaced with MBP. No shifted complexes are observed for any of the nucleic acids tested, including TAR RNA, indicating that MBP does not mediate binding for the MBP-TRBP fusion product.

3.3.2 TRBP asymmetry sensing in siRNAs

It has been hypothesized that TRBP senses siRNA asymmetry and therefore contributes to guide strand selection in RISC. In order to investigate this putative function of TRBP, we prepared three siRNAs with distinct thermodynamic features. Two of the siRNAs are known to load into *Drosophila* RISC based on their strand thermodynamics (Schwarz et al., 2003) and are consistent with $\Delta\Delta G$ energetics calculated with existing methods (Hutvagner, 2005) using current nearest-neighbor parameters and a terminal A:U penalty of 0.5 kcal/mol (Mathews et al., 1999) (see also Appendices – Materials and Methods for Ch. 3). These sequences correspond to pp-luciferase (pp-luc) (asymmetric, 'A'; $\Delta\Delta G = (\Delta G \text{ for the AS } 5\text{'-end} - \Delta G \text{ for the SS } 5\text{'-end}) = 0.5 \text{ kcal/mol}$) and human Cu, Zn-superoxide dismutase (sod1) (symmetric, 'S'; $\Delta\Delta G = 0.0 \text{ kcal/mol}$). The third sequence targets EGFP and was found to be moderately active (~60% knockdown at 33 nM concentration) from our previous work in human HeLa and HepG2 cells but is predicted to be essentially symmetric (EGFP, 'G'; $\Delta\Delta G = -0.1 \text{ kcal/mol}$) (Figure 3-5A) (Gredell et al., 2008). Position 20 of each strand was chemically modified with a 4-thiouracil to allow for position-specific photocrosslinking (Sontheimer, 1994), similar to the methods used for characterizing the *Drosophila* proteins using 5-iodouracil modified siRNAs (Tomari et al., 2004). For each of these sequences, a total of six different siRNAs were created by alternately hot ($[\gamma\text{-}^{32}\text{P}]\text{-ATP}$) or cold (ATP) end-labeling the chemically modified or unmodified single-stranded RNAs, followed by annealing with the corresponding complementary strand (Figure 3-5B, a-f). The relative crosslinking efficiency of each siRNA was then determined using recombinant TRBP in a

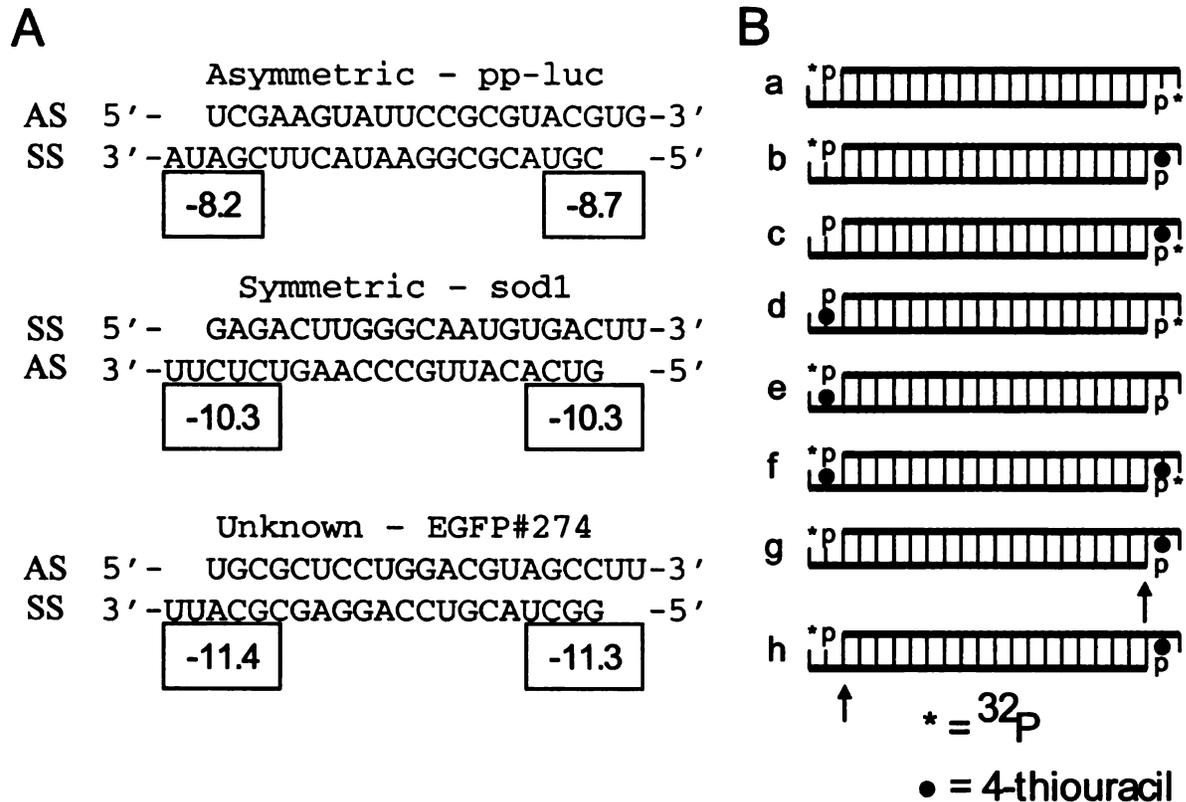


Figure 3-5 Nucleic acids used.

(A) Three siRNA sequences were used in this study. Two have been documented previously for their asymmetric (pp-luc; referred to in the text as sequence 'A') or symmetric (sod1; sequence 'S') silencing activities and recognition by Dicer-2/R2D2 in *Drosophila* (Schwarz et al., 2003; Tomari et al., 2004). The third was previously used for EGFP silencing and is expected to be slightly asymmetric based on silencing activity (Gredell et al., 2008) and $\Delta\Delta G$ calculation with mfold (Zuker, 2003) (EGFP; sequence 'G'). AS denotes the antisense strand and SS denotes the sense strand. Boxed numbers indicate ΔG values (in kcal/mol) calculated from the four terminal nearest-neighbor base-pairs, including the A:U penalty and 3'-overhang contribution (single base stacking energy). (B) A series of 8 different siRNAs were created for each sequence as described in the text Appendices – Materials and Methods for Ch. 3. The strand location in (A) matches the strand location in (B). * denotes strands 5'-labeled with ³²P-ATP; ● denotes 4-thiouracil modification at position 20; arrows indicate positions destabilized by sequence mismatches on the bottom-strand.

denaturing gel shift assay (Figure 3-6A and D). TRBP showed stronger crosslinking to the asymmetric 'b' siRNA (Ab) relative to the 'd' (Ad) structure (Figure 3-6A and D; compare Ab to Ad; $p = 1 \times 10^{-5}$) and similarly for the EGFP siRNA (Figure 3-6A and D; compare Gb to Gd; $p = 2 \times 10^{-6}$). However, almost equal crosslinking was detected for the symmetric siRNA Sb and Sd forms (Figure 3-6A and D; compare Sb to Sd; $p = 0.19$). As expected, little signal was returned when radiolabel and crosslinker were not present on the same strand, showing that crosslinking is specific for the 4-thiouracil positions (Figure 3-6A and D; compare 'b', 'd', and 'f' to 'a', 'c', and 'e'; $p < 2 \times 10^{-4}$ for all pair-wise comparisons). As further confirmation, for all three siRNAs, the crosslinking of 'f' was approximately the sum of 'b' and 'd'. Since all siRNAs were bound equally in a native gel shift assay (Figure 3-6B) and were comparably loaded (Figure 3-6C), we conclude that recombinant TRBP preferentially crosslinks to the more stably hybridized end of the siRNA duplex. Similar trends were observed using HepG2 cytoplasmic extracts (Figure 3-7), confirming that this behavior is not purely an *in vitro* artifact and that TRBP can detect siRNA asymmetry in the presence of other proteins, both competing and not.

It is critical to note that the crosslinking pattern observed for the EGFP targeting siRNA was not as expected from the $\Delta\Delta G$ calculation, with the asymmetry predicting a slight bias in favor of 'Gd' vs. 'Gb'. We observed that both the 'A' and 'G' siRNAs contained an A:U at the AS 5'-end and a G:C at the SS 5'-end, whereas the 'S' siRNA contained a G:C at both ends. Thus, bias based solely on the first base/base-pair on each end of the siRNAs would be sufficient to explain the TRBP crosslinking. A similar

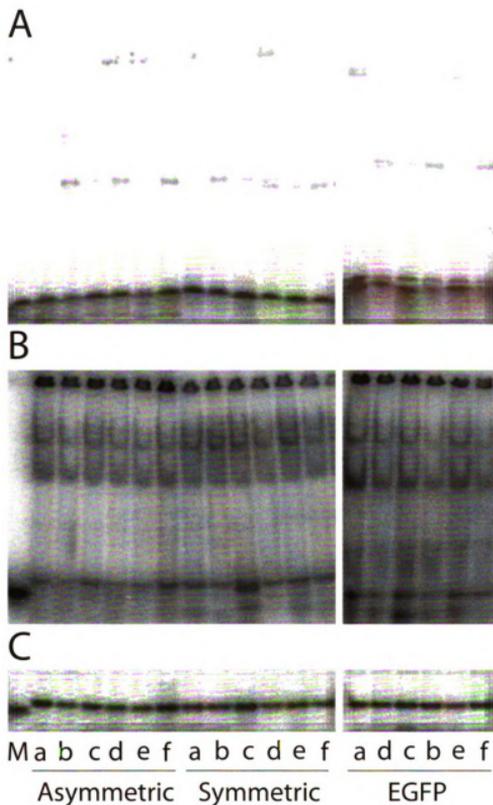
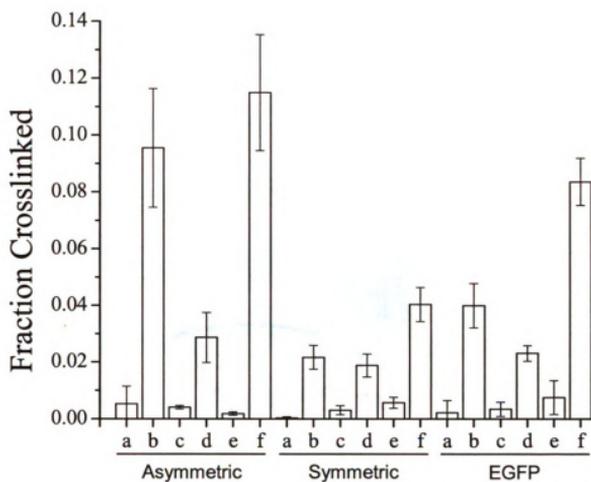


Figure 3-6 TRBP asymmetry sensing.

The siRNAs in Figure 3-5 were tested by gel shift assay with recombinant TRBP for (A) crosslinking, (B) native binding, or (C) native loading control, as described in the text and Appendices – Materials and Methods for Ch. 3. (D) The fraction of siRNA crosslinked by TRBP was quantified within each lane (fraction crosslinked = crosslinked signal / (crosslinked signal + uncrosslinked signal)). Values are average \pm standard deviation; $n \geq 5$ for 'A' and 'G'; $n \geq 3$ for 'S'. M denotes single-stranded siRNA used as denaturing size marker.

D**Figure 3-6 (cont'd).**

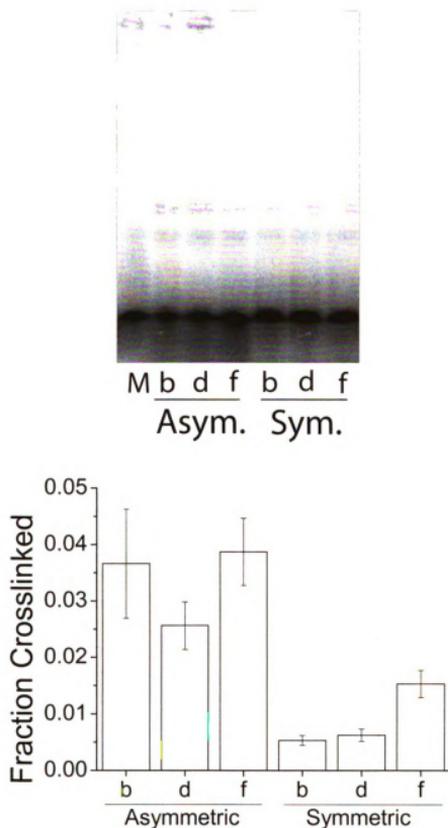


Figure 3-7 Asymmetry sensing by human cell extract.

Cytoplasmic extracts from human HepG2 cells (~19 $\mu\text{g}/\text{lane}$) were incubated with siRNA and crosslinked as in Figure 3-6. Quantification was also as described. Values are average \pm standard deviation; $n \geq 5$.

conclusion was recently reached from a computational analysis, finding that one nearest-neighbor parameter at each end of the siRNA gave the best predictive value for silencing activity (Lu and Mathews, 2007). We pursued this idea further through our own analysis of a literature dataset (Shabalina et al., 2006), as we had previously used the dataset when investigating mRNA target structure effects (Gredell et al., 2008). We sorted the tabulated siRNA activities according to the nucleotide at the 5'-end of the AS relative to the nucleotide at the 5'-end of the SS. The most active siRNAs tend to have a U at the 5'-end of the AS and a G at the 5'-end of the SS (Figure 3-8 and Figure 3-9). Generally, the order of nucleotide preference was $U > A > G > C$, for the 5'- nucleotide on the AS, and $G > C > A > U$, for the 5'- nucleotide on the SS. These observations agree completely with those made earlier by Reynolds et al (Reynolds et al., 2004). This is consistent with what would be required for a bias in stability but also implies a role for sequence, independent of stability. For instance, a U:A combination (not base-pair) is statistically more likely to be a good silencer than an A:U combination (Figure 3-8 and Figure 3-9). More work will be needed to evaluate inconsistencies with nearest-neighbor calculations and the impact of overhangs on such calculations.

3.3.3 Altering siRNA asymmetry via mismatches

Since the first and last base-pairs in the siRNA seemed to be critical for establishing siRNA asymmetry, we next considered the effect the introduction of sequence mismatches would have on TRBP asymmetry sensing (analogous to (Tomari et al., 2004) for R2D2). Single nucleotide mismatches were introduced at either position 1 or 19 of the bottom-strand of each siRNA (Figure 3-5; compare 'g' and 'h' to 'b'). The bottom-strands were cold labeled and then annealed to the hot labeled, chemically-

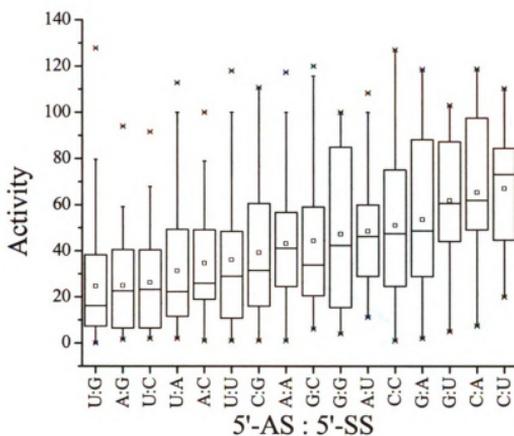


Figure 3-8 Computational analysis of first nucleotide influence on siRNA activity. siRNAs from the literature (Shabalina et al., 2006) were sorted according to the nucleotides at both the AS 5'-end and the SS 5'-end (5'-AS:5'-SS) and plotted in order of decreasing mean silencing activity. The small box indicates the mean; the center line, the median; the lower and upper box lines, the 25th and 75th percentiles; the outer bars, the 95th percentile; and the points, outliers.

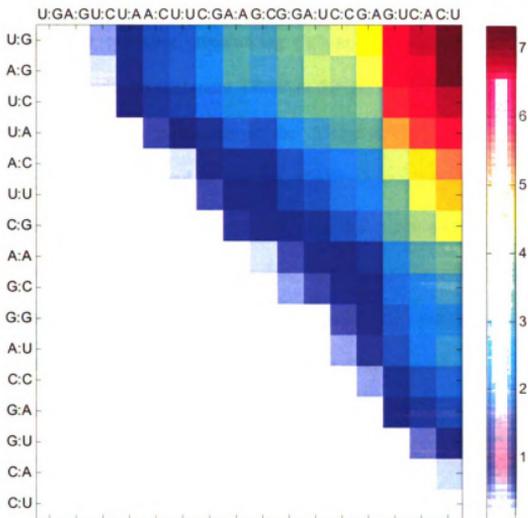


Figure 3-9 Statistical analysis for pair-wise comparison between first nucleotide on either end of the siRNAs.

A t-test was performed on every possible combination of nucleotide pairings (AS 5'-end : SS 5'-end) from Figure 3-8. The scale is based on the t-value reported by the t-test. Roughly, for the sample sizes in the dataset, a t-value > 2.0 indicates significance at the $p < 0.05$ level.

modified top-strand. The mismatched ends should be destabilized, reducing TRBP binding and either reducing crosslinking ('g') or increasing crosslinking ('h'). Each siRNA was tested with recombinant TRBP in the gel shift assays (Figure 3-10). For the asymmetric sequence, a mismatch at the 5'-end of the SS did result in reduced crosslinking (Figure 3-10A and D; compare Ag to Ab; $p = 0.008$). However, a mismatch at the 5'-end did not significantly enhance crosslinking (Figure 3-10D; compare Ah to Ab; $p = 0.49$). This may reflect the fact that this asymmetric sequence already has strongly biased TRBP crosslinking, so increasing the stability bias cannot enhance binding to the more-stable end. However, for the symmetric and EGFP sequences, neither mismatch resulted in any statistically significant change in crosslinking (Figure 3-10D; compare Sg and Sh to Sb, $p > 0.31$; and Gg and Gh to Gb, $p > 0.28$), which may be due to the ~2-fold lower crosslinking compared to the asymmetric sequence despite the fact that all native binding experiments showed nearly complete binding.

As we saw similar binding but different crosslinking among the three siRNAs, we wanted to revisit TRBP binding and crosslinking of DNA/RNA hybrids (Figure 3-10). For all three of the sequences tested, none was appreciably bound (Figure 3-2, lanes 9-12, where the sequences of the heteroduplexes tested - for binding only - correspond to the EGFP sequence; also Figure 3-10B, 'i' and 'j', where 'i' and 'j' correspond to 'b' and 'd' in Figure 3-5, with the bottom- or top-strand of the siRNAs replaced with DNA, respectively) or crosslinked by TRBP (Figure 3-10A and D, 'i' and 'j'). Since it has also been suggested recently that 5'-O-methyl modifications can alter guide strand selection and the propensity for off-target silencing (Chen et al., 2008), we created EGFP siRNAs with the 5'-O-methyl modification on either end and the 4-thiouracil-modified

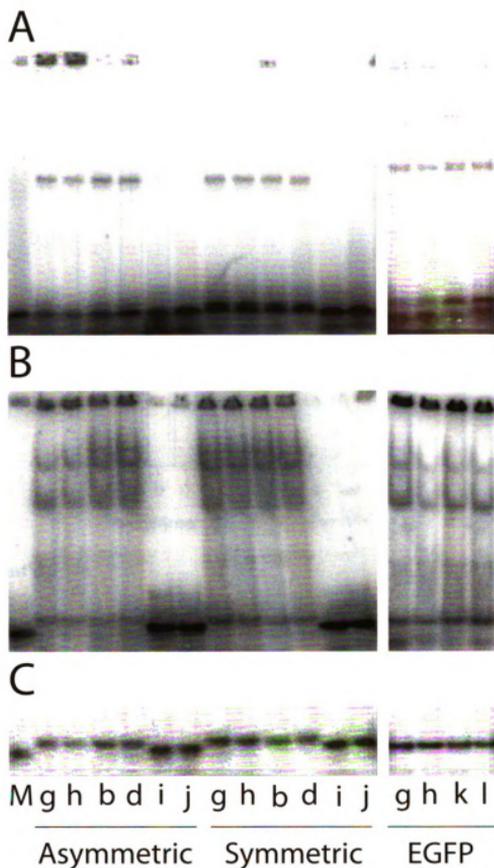


Figure 3-10 Mismatched siRNAs alter TRBP asymmetry.

The siRNAs were tested by gel shift assay for (A) crosslinking, (B) native binding, or (C) native loading control, and (D) the fraction of siRNA crosslinked by TRBP was quantified within each lane, as in Figure 3-6. Values are average \pm standard deviation; $n \geq 3$. M denotes single-stranded siRNA used as denaturing size marker.

D

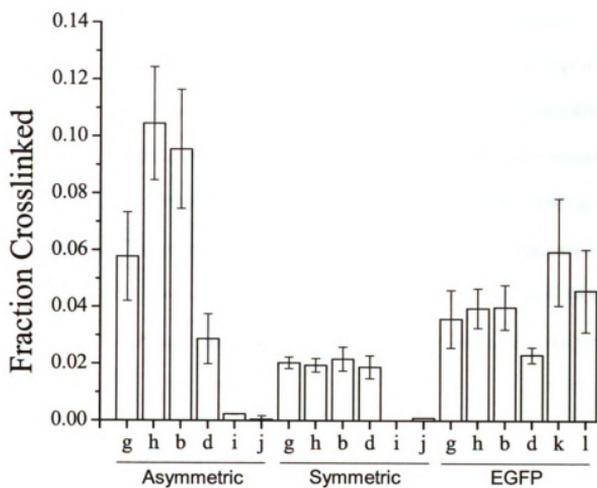


Figure 3-10 (cont'd).

complementary strand being hot labeled (Gk with the bottom-strand containing the 5'-O-methyl corresponds to Gg because the 5'-O-methyl groups are only available on T nucleotides and consequently introduced a C:T mismatch at that end; Gl corresponds to Gd with the top-strand modified). We expected binding and therefore crosslinking in both cases to decrease due to the proximity of the 5'-O-methyl group and the 4-thiouracil modification. However, the methylation had no apparent effect on TRBP binding, and instead increased the overall amount of crosslinking (Figure 3-10A and D; compare Gk to Gg, $p = 0.026$, and Gl to Gd, $p = 0.012$). Thus, precluding the interaction with TRBP is likely not the means by which these modifications prevent incorporation of these strands into RISC.

3.3.4 TRBP recognition of single-stranded small RNAs

Given that TRBP can sense siRNA asymmetry, and despite its apparent inability to bind short ssRNA (Figure 3-2), we also wanted to consider whether any other short ssRNA sequences could be bound or crosslinked by TRBP. We repeated the gel shift assay with the individually hot labeled, chemically modified ssRNAs from each of the three siRNA sequences, at increased TRBP concentration (~3-fold higher; 1200 nM) (Figure 3-11). Intriguingly, the strand favorably crosslinked when part of the asymmetric siRNAs (Figure 3-6A and D, 'b' for 'A' and 'G') was also preferentially crosslinked here as a single-strand (Figure 3-11A and D; compare AS to SS for 'A' and 'G'), whereas the single-strands corresponding to the symmetric siRNA were crosslinked more evenly (Figure 3-11A and D; compare AS to SS for 'S'). Curiously, the native binding pattern did not appear to reflect that of the crosslinking gel (Figure 3-11B), which might be expected for a relatively low affinity interaction that could be altered by the conditions

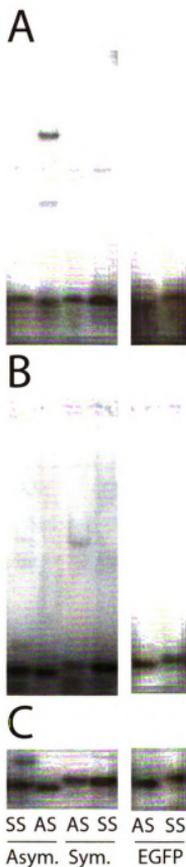


Figure 3-11 TRBP recognition of short single-stranded RNAs.

The ssRNAs from each siRNA duplex were tested by gel shift assay for (A) crosslinking, (B) native binding, or (C) native loading control, and (D) the fraction of siRNA crosslinked by TRBP was quantified within each lane, as in Figure 3-6. Values are average \pm standard deviation; $n \geq 4$. The values listed below the graph are the ratios between the AS crosslinking and the SS crosslinking.

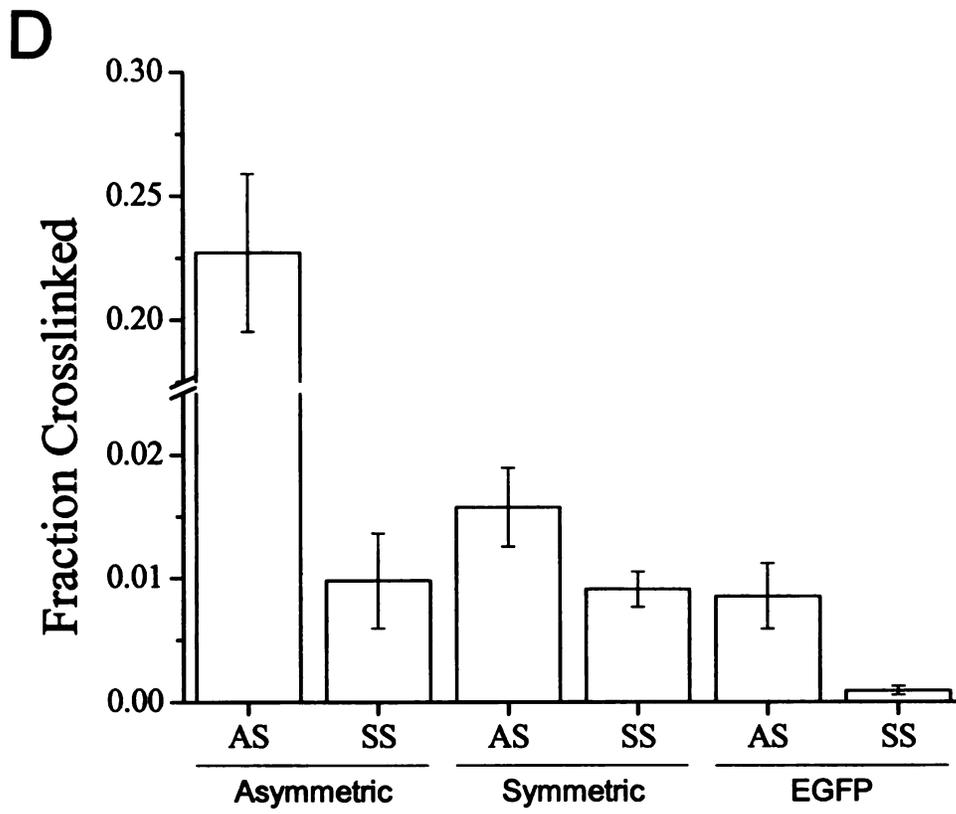


Figure 3-11 (cont'd).

for electrophoresis. While this supports our contention that terminal sequence matters in TRBP-RNA interactions, further exploration will certainly be required to explain this result and to ascertain the influence of 5'-end versus 3'-end effects.

3.4 Discussion

RNA interference relies on guide strand complementarity to identify the target for inhibition. Although the manner by which the guide strand is selected in humans has been elusive, it is known to be accomplished collectively by the proteins of the RLC: Dicer, TRBP, and Ago2. Our results show for the first time that human TRBP alone senses the asymmetry of an siRNA duplex *in vitro*, a critical step in the guide strand selection process.

An important step in RLC formation is the initial binding of the double-stranded siRNA duplex. Dicer and Ago2 appear to be less capable of binding siRNA alone (Kini and Walton, 2007; Rivas et al., 2005), while TRBP shows a relatively higher affinity ($K_D \sim 1$ -100nM; data not shown and (Parker et al., 2008)). When in solution, TRBP homodimerizes (Kok et al., 2007; MacRae et al., 2008), which may contribute to TRBP forming multimeric complexes on dsRNA even as short as 21 base-pairs (e.g., siRNA; Figure 3-2 and Figure 3-3) in an uncooperative manner (Parker et al., 2008). This perhaps indicates a role different from dsRBPs that show improved binding with dsRNA length, such as *C. elegans* RDE-4 (Parker et al., 2008). Consistent with what has been observed for other dsRBPs and with dsRBD function in general (Bevilacqua and Cech, 1996; Saunders and Barber, 2003), TRBP was unable to bind double-stranded siDNA or RNA/DNA hybrids (Figure 3-2 and 3-10). Recent evidence shows that some DNA

substitutions are fairly well tolerated in siRNAs for gene silencing, particularly when present at the 5'-end of the guide strand (Ui-Tei et al., 2008). Substitutions at the 3'-end, however, could disrupt silencing, presumably by affecting the TRBP binding site. These results are in agreement with our observations (Figure 3-2 and 3-10), and it would be interesting to see how shorter DNA segments spread throughout the siRNA would impact binding and crosslinking by TRBP.

Though all siRNAs appear to be bound by TRBP with similar high affinity, and apparently independent of sequence (Figure 3-6B and 2-10B and data not shown), in our hands, TRBP showed a clear bias in crosslinking, seemingly driven by both sequence and strand thermodynamics (Figure 3-6). TRBP recognized two siRNAs asymmetrically ('A' and 'G') and one symmetrically ('S') in our crosslinking gel shift assay (Figure 3-6A and D). For two of the siRNAs tested, our results are consistent with those previously shown in *Drosophila* RISC loading (Schwarz et al., 2003) and in crosslinking by Dicer-2/R2D2 (Tomari et al., 2004), as well as when using standard techniques to calculate relative thermodynamic asymmetry (Hutvagner, 2005) (see also Appendices – Materials and Methods for Ch. 3) (Figure 3-5A). Interestingly, the third sequence ('G') is predicted to be only minimally asymmetric ($\Delta\Delta G = -0.1$ kcal/mol), but favoring the opposite end from what was experimentally determined (Figure 3-5 and 3-6, Gb vs. Gd). However, one might anticipate the experimentally observed asymmetry based on the sequence of the 'G' duplex (Figure 3-5). The end favorably crosslinked in Gb contains a G:C terminal base-pair (at position 19 of the AS), whereas the opposite end (pos. 1) is an A:U pair. This is also true of the 'A' siRNA, but not for the 'S' siRNA (there is a G:C pairing at both ends). Thus, it would appear that the single base-pair at either end is sufficient for predicting

asymmetry (Figure 3-8 and Figure 3-9). A recent computational analysis reached a similar conclusion that the best predictive results were found when asymmetry only included the first base-pair of each end (Lu and Mathews, 2007). Furthermore, when Schwarz et al., replaced either of the end G:C base-pairs in the S sequence with an A:U pair, the siRNA became asymmetric (Schwarz et al., 2003). It may also be that the relative thermodynamics are insufficient to predict asymmetry and instead the absolute energetics may also be important, as has been previously suggested (Schwarz et al., 2003).

In *Drosophila*, it was shown that the RLC (Dicer-2/R2D2) initiates unwinding of the siRNA, containing a small amount of ssRNA relative to siRNA (Tomari et al., 2004). We wondered if TRBP might perform a similar function in humans as it has been shown to destabilize the TAR RNA secondary structure by base unstacking in the region of the upper-stem/loop (Erard et al., 1998). Although exceedingly high, non-physiological TRBP concentrations were required, we detected some low-level ssRNA binding activity (Figure 3-11A). This low affinity interaction with ssRNA echoes what our group and others have previously shown for Dicer (Kini and Walton, 2007; Lima et al., 2009). More surprisingly, perhaps, TRBP preferentially crosslinked to the ssRNA in a fashion that reflected the biased loading of the strands into active RISC (Figure 3-11B; compare AS to corresponding SS for each sequence) and the overall level of duplex crosslinking ('A' > 'G' > 'S'). We examined if ssRNA secondary structure was responsible for differences in binding/crosslinking, but no patterns were discernible for structures predicted by mfold (Zuker, 2003) (data not shown). Likewise, we cannot rule out that these effects were merely due to non-specific crosslinking of ssRNA and unbound

protein. However, the fact that the ssRNA crosslinking was consistent with that detected for the duplex seems more than coincidental and will require further testing before a definitive conclusion can be reached.

Accurately predicting which siRNA strand will be loaded into active RISC, whether based on the individual end base-pairs, the relative energy differential, or some other metric, is essential for optimizing siRNA activity. Current algorithms tend to weigh asymmetry heavily (Lu and Mathews, 2007; Shao et al., 2007). There are several reports that use mismatches to ensure the desired asymmetry. Interestingly, when we introduced a single nucleotide mismatch in the asymmetric pp-luc siRNA at the end preferentially crosslinked by TRBP, we observed a significant decrease in crosslinking (Figure 3-10, Ag vs. Ab), although the overall asymmetry still favored that end. The corresponding mismatch at the opposite end did not significantly increase crosslinking (Figure 3-10, Ah vs. Ab), perhaps indicating that there is a maximum bias beyond which TRBP crosslinking is not enhanced. Unexpectedly, for the 'S' and 'G' siRNAs, the corresponding mismatches had no effect on TRBP crosslinking (Figure 3-10). It is unclear why this was the case since the mismatches in S were previously shown to alter its loading in *Drosophila* (Schwarz et al., 2003; Tomari et al., 2004), but it may again reflect the importance of sequence and relative versus absolute strand thermodynamics. It could also be that more than one mismatch is needed to alter the asymmetry, as was recently shown (Geng and Ding, 2008). Other work by our group using human cell extracts found that a single nucleotide mismatch at the guide strand 5'-end of an siRNA can cause reduced TRBP binding, with a concomitant reduction in silencing efficacy (H. K. Kini and S. P. Walton, unpublished results). Or, it may be a difference in the function

of the human and *Drosophila* systems. TRBP alone may not entirely recapitulate in humans the function of Dicer-2/R2D2 in *Drosophila*, requiring also Dicer and/or other components. This is also supported by other work from our group indicating that Dicer binding to an siRNA is improved by incorporation of a terminal mismatch (H. K. Kini and S. P. Walton, unpublished results).

Several chemical modifications have previously been tested with varying degrees of influence on activity (Chiu and Rana, 2003; Sipa et al., 2007), but it is unknown to what extent the changes in activity are results of changes in TRBP binding and asymmetry sensing. We tested one such modification, 5'-O-methylation, shown to control guide strand selection and eliminate off-target silencing by the modified strand in humans (Chen et al., 2008). Presumably, the 5'-O-methyl group prevents phosphorylation by hClp (Weitzer and Martinez, 2007) and subsequently its incorporation into RISC. In our assay, the 5'-O-methylated siRNAs were more effectively crosslinked than the unmodified versions (Figure 3-10A and D), with no change in overall TRBP binding (Figure 3-10B). This suggests that 5'-phosphorylation is not essential for TRBP asymmetry sensing and that the cellular effects on activity or off-target silencing are a result of additional factors associated with RISC, such as Dicer for which it was shown the 5'-O-methyl modification can disrupt binding (Pellino et al., 2005). It would be interesting to see how other modifications that block/enhance RISC activity, such as A to I RNA editing of pri-miRNA (Kawahara et al., 2007), would affect TRBP binding and crosslinking as it has been shown for at least one dsRBP (PKR) that modifications that abrogate activity do not necessarily have an effect on dsRNA binding (Nallagatla and Bevilacqua, 2008).

The interaction(s) and positioning between Dicer/TRBP/Ago2 are critical for proper loading of the siRNA guide strand onto Ago2 and for functional RNAi activity. We cannot rule out the contributions of these or other associated proteins, but our results here show that TRBP alone can bind siRNAs and sense thermodynamic asymmetry. Although Dicer and Ago2 have not been shown to perform that function in RISC, Dicer may assist in asymmetry sensing, as it appears to do in *Drosophila* (Tomari et al., 2004). Quite possibly, Dicer also assists in siRNA unwinding, perhaps through its helicase domain which was recently shown to be important for processing thermodynamically unstable shRNAs (Soifer et al., 2008). Furthermore, the interaction between the RNase III domain of Dicer and the PIWI domain of Argonautes (Tahbaz et al., 2004) could facilitate loading of the 5'-end of the guide strand onto Ago2 (Rivas et al., 2005). In *Drosophila*, Ago2 is necessary for completely unwinding the siRNA (Okamura et al., 2004), which may suggest that in humans Dicer/TRBP can facilitate that step of the processing. Interestingly, a reduction in either TRBP, Dicer, PACT, or RHA seems to reduce si/miRNA processing and RISC activity, an effect that may or may not be due to destabilization of other RISC proteins (Chendrimada et al., 2005; Haase et al., 2005; Kok et al., 2007; Lee et al., 2006; Maniataki and Mourelatos, 2005; Robb and Rana, 2007). More experiments will be needed to detail the collective contributions of these proteins to RNAi and to see if TRBP is the only RISC protein that senses asymmetry.

CHAPTER 4 POLYMERIC NANOPARTICLES FOR siRNA DELIVERY

4.1 Abstract

The selection of siRNA sequences is clearly an important factor contributing to the effectiveness of RNAi applications. But no matter how active a particular siRNA is during initial trials in cell culture, a major determinant of their eventual widespread therapeutic use *in vivo* is delivery. A variety of methods are being developed to accomplish both systemic and localized delivery and typically utilize either lipid- or polymer-based preparations. Despite the success of some of these methods to deliver *in vivo*, there remains a general lack of information regarding the lipid/polymer contributions to nucleic acid binding, nuclease protection, cellular uptake, toxicity, and release of the siRNA, primarily due to vast differences in the systems being developed. Furthermore, the existence of diverse experimental data published thus far can at least in part be attributed to the difficulty of modifying many of the lipids/polymers.

The work described here was aimed at developing a novel biocompatible and biodegradable cationic polymer nanoparticle (NP) capable of delivering siRNA into cells grown *in vitro*, with potential use later for *in vivo* applications. The polymer developed by our collaborators was readily modified by standard “click” chemistry methods, which allowed for a systematic study of variables contributing to siRNA binding. Strong siRNA binding was found to require an NP containing a combination of primary and secondary amine groups to facilitate electrostatic interactions. Inclusion of alkyl chains enhanced binding, perhaps by causing vesicle-like formation. Consistent with its known charge shielding effects, polyethylene glycol (PEG) could be included to reduce overall

binding affinity, and is expected to be important for nuclease protection. Ultimately, the NPs were capable of binding and then delivering siRNAs to cells, but were unable to initiate RNAi. Accordingly, subsequent studies will be needed to investigate how alterations to the polymer affect cellular uptake and siRNA release without greatly affecting NP/siRNA complex formation so as to maximize the silencing potential of this siRNA delivery platform.

Author contributions:

The work described in this chapter was performed in collaboration with Professor Gregory Baker from the Michigan State University Department of Chemistry, his graduate student Erin Vogel, and Professors Christina Chan and S. Patrick Walton from the Michigan State University Department of Chemical Engineering and Materials Science.

We all contributed to the project conception, experimental design, and data analysis. Erin was responsible for synthesis of the polymers tested and for characterization of NP/nucleic acid complexes by transmission electron microscopy and dynamic light scattering. I performed the gel shift assays and curve fit analyses to determine NP binding parameters, with the assistance of several undergraduate students (Sophie Carrell, Dan Desantis, and Jorge Fontes), as well as all cell treatment experiments. The chapter itself was written entirely by me with minor editing by Professor Walton.

4.2 Introduction

The therapeutic potential of siRNAs greatly relies on being able to overcome the natural barrier of human cells to foreign material, in particular highly charged species like nucleic acids. Nucleic acids do not freely diffuse through the cellular membrane, are rapidly degraded by nucleases, bound by other factors found in the blood, and cleared by the kidneys (Zhang et al., 2007). The intentional delivery of dsRNA into human cells also presents a unique challenge because of the potential for stimulating an immune response; viral methods for gene therapy can overcome some of these limitations, yet they still suffer from a number of safety issues. Instead, much effort is being spent developing non-viral delivery methods for siRNAs so that their therapeutic potential can be realized.

As an alternative, both lipid and polymer based reagents are being pursued as delivery vehicles that can be applied systemically, with the potential for localized targeting, and that may even allow transport through the blood brain barrier for treatment of neurological disorders (Pardridge, 2007). However, lipid-based delivery reagents (reviewed in (Tseng et al., 2009)) present their own drawbacks, such as potentially high *in vivo* toxicity (Zhang et al., 2007) and a relatively limited chemical and structural diversity. Therefore, considerably more effort has been invested in creating polymeric vehicles for nucleic acid delivery, in general, and for delivery of siRNAs, specifically (reviewed in (de Martimprey et al., 2009)).

Wide varieties of polymeric vehicles have been synthesized by various chemical means, but, as with viral methods and lipid carriers, are ultimately restricted based on

their biostability and biocompatibility. For many of the polymers traditionally used for nucleic acid delivery, such as PEI and PLL, biostability and biocompatibility can be improved by attaching groups like PEG (for charge shielding and to facilitate membrane interactions) or other biomolecules (e.g., cholesterol) through standard chemical means (Kircheis et al., 2001; Swami et al., 2007; Urban-Klein et al., 2005; Wolff and Rozema, 2008). However, the steps involved in modifying these polymers can require diverse reaction conditions and solvents.

A new method for post-polymerization modification was proposed in 1999 by Dr. Barry Sharpless' group that was aimed at mimicking biology by using a series of highly specific, high yield reactions that can be performed under relatively benign conditions and require little in the way of product isolation (Kolb et al., 2001; Sharpless and Kolb, 1999). Today, these reactions, known as "click" chemistry, have become a powerful tool for pharmaceutical applications due to the ease with which they are performed. Perhaps the most prominent reaction meeting the "click" criteria is the Huisgen 1,3-dipolar cycloaddition of azides and terminal alkynes (reviewed in (Hein et al., 2008)). This reaction has been implemented by our collaborators (Jiang et al., 2008) in the development of a novel polymer system that resembles poly(lactic-co-glycolic acid) (PLGA), another polymer frequently used for nucleic acid delivery (de Martimprey et al., 2009). Both PLGA, and this polymer, poly(propargylglycolide) (PPGL), are biocompatible and biodegradable (into lactic acid derivatives), thereby lending themselves to a biological setting. However, unlike PLGA, PPGL can be readily modified by "click" cycloaddition reactions (through the formation of 1,2,3-triazoles) to

create new materials (Figure 4-1A) with potentially enhanced properties (Jiang et al., 2008).

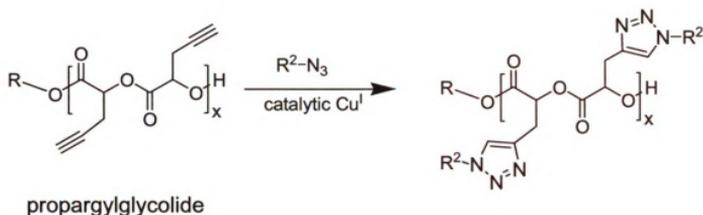
These characteristics make PPGL an ideal candidate for the encapsulation of small molecules, including siRNA, for their delivery as therapeutics *in vivo*. Furthermore, because the azide group is relatively easily added to many of the moieties already being included in other complex polymer systems (Figure 4-1B), a family of related vehicles can be created that incorporates nearly limitless modifications. The final structure when the azide groups are added to the propargyl backbone is roughly brush-like, with the azide groups forming the bristles of the brush (Figure 4-1C). Accordingly, with this polymer backbone it is possible to systematically investigate the effect various moieties have on siRNA vehicle design.

4.3 Results and Discussion

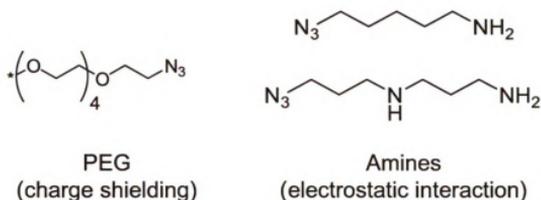
Our approach was to begin developing the PPGL polymer system developed in Prof. Baker's laboratory (Jiang et al., 2008) to deliver siRNA to cells grown in culture by first systematically studying the impact of functional group additions to the polymer backbone, in particular the influence the various groups had on NP binding of siRNA. Description of the polymer syntheses and other experimental details can be found in Appendices – Materials and Methods for Ch.4. Future studies to be performed by others will aim to characterize and engineer other important parameters, such as nuclease protection, cellular uptake, cytoplasmic release of the siRNA (endosomal escape), and activity in RNAi.

We first determined whether a preliminary version of the polymer (unmodified PPGL) could enter the cells and if so, what was its cellular distribution. A rhodamine

A



B



C

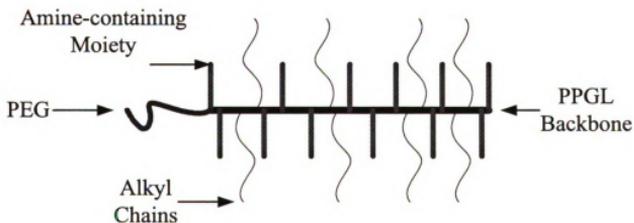


Figure 4-1 Nanoparticle synthesis details.

(A) Nanoparticles were created by adding functional groups (B) to the poly(propargyl glycolide) (PPGL) backbone. Amine-containing moieties or polyethylene glycol (PEG) provide positive charge and charge shielding, respectively. Additional types of groups can be added, as listed in Table 4-1 and shown in Figure 4-4. (C) A sample nanoparticle is depicted. Attached to the polymer backbone is a PEG tail (left end), and amine-containing moieties and alkyl chains dispersed randomly along the backbone. The backbone and sidechains are flexible and relatively free to organize into NPs in solution.

fluorescent group was attached to the polymer and the solution was administered to human HeLa cells for 24 hrs. Fluorescence microscopy confirmed that the cells appeared healthy and were indeed internalizing the polymer into the cytoplasm (Figure 4-2A).

Most polymers typically pursued for siRNA delivery form complexes with the siRNA through electrostatic interactions with the anionic nucleic acid backbone. Frequently, the siRNA and vehicle are simply mixed in an aqueous buffer and allowed to self-assemble to form condensed complexes. We therefore considered whether the NP in question could form a stable complex with nucleic acid as that is the first step in delivery. Rather than use siRNA to establish binding characteristics, we instead assessed binding with a dsDNA molecule comparable in length to an siRNA and containing 2 nt 3'-overhangs (siDNA). The charges for the two types of nucleic acid are the same, albeit with slightly different organization along the backbone because of the difference in helix geometry, but the siDNA is approximately 1/10th or less the cost of the siRNA. The polymer and siDNA were mixed in water at room temperature for 15 min and then subjected to agarose gel electrophoresis to separate bound and unbound nucleic acid (gel shift assay). No complex formation was observed for this NP (data not shown). However, shifted complexes were visualized with this method using several of the polymers discussed above, including LPEI, BPEI, and the lipid solution Lipofectamine 2000 (LF2k; Invitrogen), indicating that this assay could be used to screen alternate polymer chemistries (data not shown). The gel shift assay has also been applied by others for testing polymer/plasmid DNA or polymer/siRNA binding characteristics with similar success (Bartlett and Davis, 2007b; Jeong and Park, 2002; Mann et al., 2008).

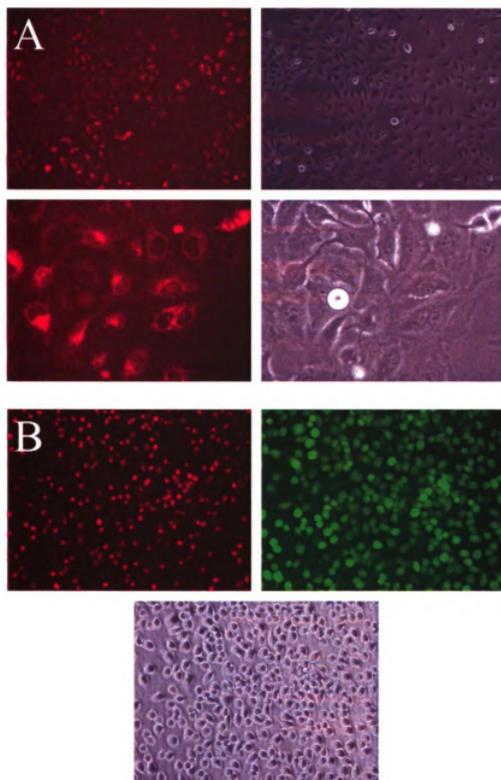


Figure 4.2 Visualization of nanoparticle delivery into cell culture by fluorescence microscopy.

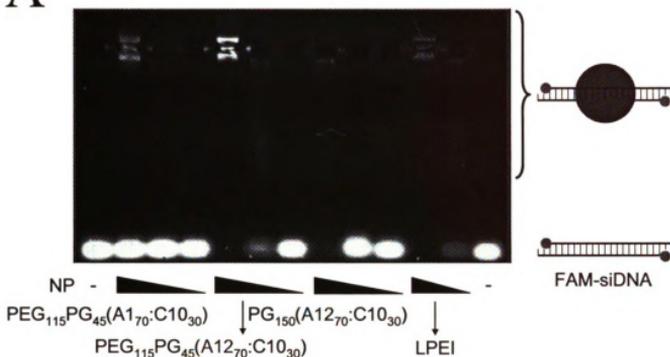
(A) Images show uptake of rhodamine tagged NP (1.6 $\mu\text{g/ml}$) into HeLa cells. Top row: 10x mag. Bottom row: 40x mag. Left column: fluorescence. Right column: phase contrast. (B) Delivery of Dy547-tagged siRNA (30 nM; red fluorescence; upper left) by untagged NP (1 mg/ml) into H1299 cells stably expressing the Enhanced Green Fluorescent Protein (EGFP; green fluorescence; upper right). The cells tolerated the NP/siRNA and remained healthy (phase contrast; bottom center). High EGFP levels indicate no silencing due to siRNA delivery. All images 10x mag.

We hypothesized that the failure of the preliminary NP to bind siDNA was due to little or no electrostatic interaction between the NP and siDNA. Subsequently, several new NPs were synthesized to investigate the effects on binding of 1) the different types and amounts of amine groups that could be added to the backbone; 2) the location and length of polyethylene glycol (PEG; for its charge shielding and membrane interaction effects); and 3) the addition of alkyl chains dispersed among the amine groups (for facilitating potential vesicle-like formation). Each NP was then tested by gel shift assay (Figure 4-3A). The fraction of siDNA bound, calculated from the fluorescence signal for the unbound siDNA, was plotted versus NP concentration (Figure 4-3B) and the resulting binding data were regressed according to the Hill equation for describing nonlinear dose-response behavior (Goutelle et al., 2008)

$$\textit{FractionBound} = \frac{[NP]^n}{K^n + [NP]^n}$$

to determine the binding constant (K , in $\mu\text{g/ml}$) and cooperativity of the binding interaction (n) for each chemistry (Table 4-1 and Figure 4-4). It is important to note that the parameters K and n of the Hill Equation have physical meaning in only limited situations (Goutelle et al., 2008; Weiss, 1997), which may or may not be applicable here. The equation is used only as a means to compare binding between different polymer chemistries without having prior knowledge of the mechanism of interaction between NP and siDNA.

A



B

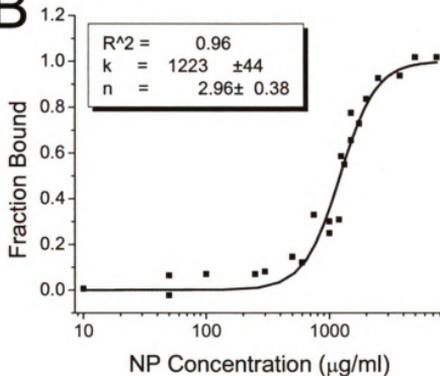


Figure 4-3 Nanoparticle complex formation with siDNA (analog of siRNA).

(A) siDNA was incubated with decreasing concentrations of NP (or control polymer, linear polyethylenimine; LPEI) for 15 min and separated by gel electrophoresis. (B) Sample fit of fraction bound versus NP concentration fitted with the Hill equation. K is the binding constant ($\mu\text{g/ml}$) and n is the cooperativity. The nomenclature corresponds to the structures described in Table 4-1 and depicted in Figures 4-1 and 4-4.

Table 4-1 Binding strength (*K*) and cooperativity (*n*) of nanoparticles for siDNA as determined by gel shift assay.

Name	Structure	Hill Equation			
		<i>K</i> ($\mu\text{g/ml}$)	+/-	<i>n</i>	+/-
VIII.7	PEG ₁₁₅ PG ₄₅ (A170C1030)	> 10,000			
VIII.8	PEG ₁₁₅ PG ₄₅ (A1270C1030)	1224	45	2.96	0.38
VIII.9	PG ₁₅₀ (A1270C1030)	776	33	1.71	0.18
VIII.13	PEG ₁₁₅ PG ₄₅ (A1290C1010)	275	15	1.21	0.09
VIII.14	PEG ₁₁₅ PG ₄₅ (A1280C1020)	236	17	1.65	0.21
VIII.15	PEG ₁₁₅ PG ₄₅ (A1260C1040)	> 10,000			
VIII.16	PEG ₁₁₅ PG ₄₅ (A1250C1050)	> 10,000			
VIII.19	PEG ₈ PG ₃₅ (A1270C1030)	70.5	3.3	2.42	0.35
VIII.21	PEG ₁₁₅ PG ₄₅ (A1290X10)	1198	104	1.27	0.17
VIII.22	PEG ₁₁₅ PG ₄₅ (A1270X30)	2639	207	0.79	0.07
VIII.27	PEG ₈ PG ₄₅ (A1290C1010)	52.5	5.0	1.43	0.19
VIII.28	PEG ₈ PG ₄₅ (A1280C1020)	31.5	3.1	1.43	0.16
VIII.29	PEG ₈ PG ₄₅ (A1270C1030)	108	21	1.27	0.36
VIII.30	PEG ₈ PG ₄₅ (A1260C1040)	119	16	1.18	0.2
VIII.43	PEG ₁₁₅ PG ₄₅ (A180C1020)	> 10,000			
VIII.44	PEG ₁₁₅ PG ₄₅ (A280C1020)	> 10,000			
VIII.45	PEG ₁₁₅ PG ₄₅ (A140A240C1020)	> 10,000			
VIII.46	PEG ₁₁₅ PG ₄₅ (A380C1020)	> 10,000			

Table 4-1 (cont'd).

Name	Structure	Hill Equation		
		K (µg/ml)	+/-	n +/-
VIII.47	PEG ₁₁₅ PG ₄₅ (A1280C1620)	63.9	1.4	2.62 0.17
VIII.48	PEG ₈ PG ₁₀ (A1280C1020)	53.2	3.7	1.38 0.13
VIII.51	PG ₅₅ (C1EO480C1020)	> 10,000		
VIII.53	PEG ₈ PG ₄₅ (A1290X ₁₀)	85.9	5.0	1.27 0.11
VIII.54	PEG ₈ PG ₄₅ (A1280X ₂₀)	484	46	0.85 0.08
VIII.55	PEG ₈ PG ₄₅ (A1270X ₃₀)	2248	270	0.55 0.04
VIII.57	PEG ₈ PG ₄₅ (A280C1020)	> 10,000		
VIII.58	PEG ₈ PG ₄₅ (A380C1020)	> 10,000		
VIII.59	PEG ₈ PG ₄₅ (A1280C1620)	13.3	1.0	1.94 0.28
VIII.65	PEG ₈ PG ₄₅ (A1280C1EO4C620)	173	15	1.06 0.09
VIII.68	PEG ₈ PG ₄₅ (A1280C8p820)	52.8	4.6	1.65 0.24
VIII.69	PEG ₈ PG ₄₅ (A1270X ₃₀)	> 10,000		
VIII.71	PEG ₁₁₅ PG ₄₅ (A1280X ₂₀)	261	22	0.97 0.09
VIII.72	PEG ₁₁₅ PG ₄₅ (A1280C1620)	154	14	1.33 0.18
VIII.73	PEG ₈ PG ₄₅ (A1275C1020G ₅)	242	20	1.68 0.21
VIII.83	PEG ₈ PG ₄₅ (A1270AP ₁₀ C1020)	203	15	1.42 0.16
VIII.84	PEG ₈ PG ₄₅ (A1280C420)	135	9	1.01 0.08
VIII.85	PEG ₁₁₅ PG ₄₅ (A1280C420)	365	36	0.90 0.09

Table 4-1 (cont'd).

Name	Structure	Hill Equation		
		K ($\mu\text{g/ml}$)	+/-	n +/-
VIII.86	PG ₄₅ (A1280C1EO3C12 ₂₀)	71.0	4.7	2.04 0.30
VIII.94	PEG ₁₁₅ PG ₄₅ (A1280X ₂₀)	1255	167	0.73 0.09
VIII.104	PEG ₈ PG ₄₅ (A12280C10 ₂₀)	13.0	1.0	1.19 0.12
VIII.112	PEG ₈ PG ₄₅ (A12250C1EO4C630X ₂₀)	28.3	3.9	0.95 0.11
VIII.116	PEG ₈ PG ₄₅ (A1270C10 ₂₀ CL ₁₀)	142	28	0.6 0.08
PLL		0.358	0.063	0.92 0.16
LPEI2.5		3.17	0.56	1.07 0.16
LPEI25		2.25	0.39	1.78 0.33
BPEI		37.8	1.8	2.64 0.27

The NPs were named according to the order in which they were synthesized. Their structures follow the general nomenclature for the groups shown in Figure 4-4 according to the following: PEG_aPG_b(A_zC_y...) where a is the number of repeat units of polyethylene glycol (PEG), b is the backbone length (number of repeat units of propargyl glycolide, PG), A is the type of amine-containing group filling z% of the available reactive (alkyne) sites, C is the type (length) of carbon-containing group filling the remaining reactive sites (y=100%-z; C=0 is denoted as X). Note that in some instances additional groups were added to the backbone, in which cases the scheme was logically extended (e.g., VIII.45). The binding constant (K), cooperativity (n), and their corresponding errors (+/-) were determined by curve fitting of gel shift data with the Hill Equation using Microcal Origin software.

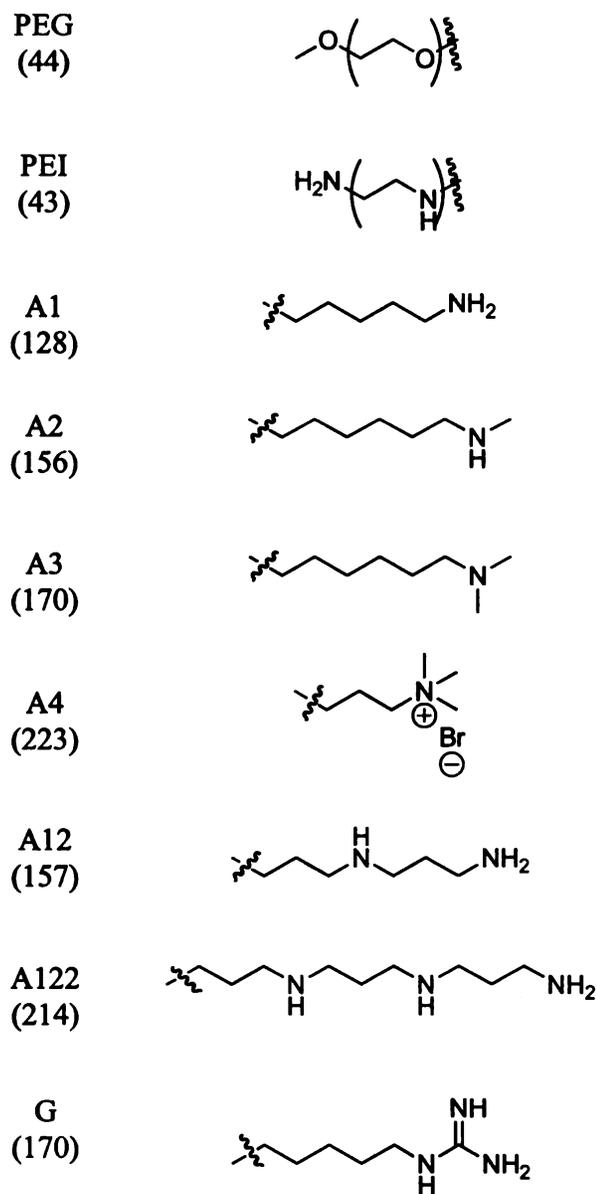


Figure 4-4 Nanoparticle functional groups added to the propargyl glycolide backbone.

The NPs listed in Table 4-1 (and data not shown) were synthesized by incorporating these groups at the 'R' position (Figure 4-1A) or by reacting with the alkyne positions (to yield the 'R²' groups). PEG = polyethyleneglycol. PEI = polyethylenimine. D = dabcyI. R = rhodamine.

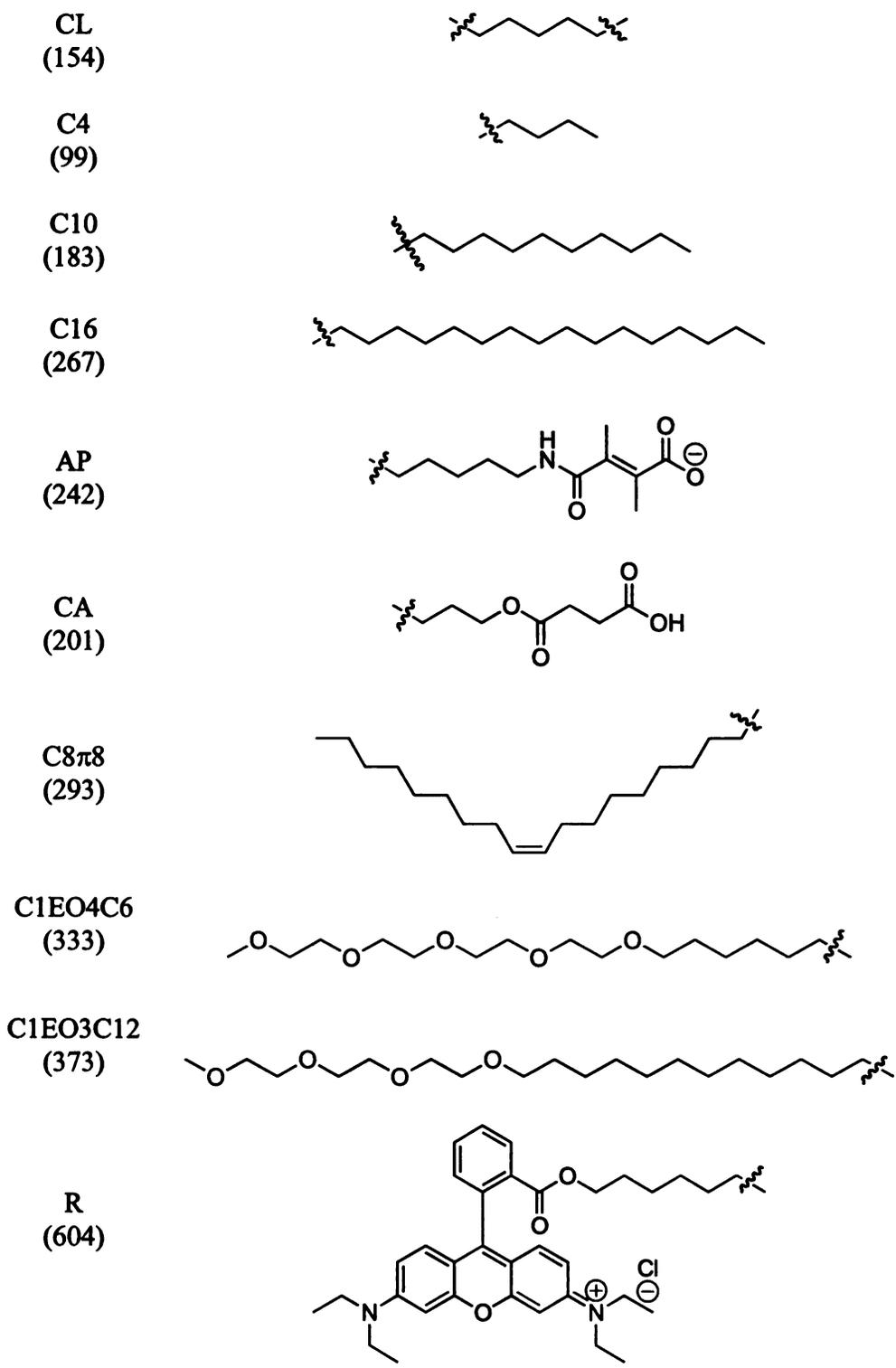


Figure 4-4 (cont'd).

It was clear from the nearly 100 different polymers tested by gel shift assay that the type and amount of amine-containing groups attached to the polymer backbone significantly affected siDNA binding characteristics (Table 4-1). NPs containing no amine groups (no A) or only primary (A1), secondary (A2), tertiary (A3), or a combination of primary and secondary amine-containing groups on separate chains (A1A2), showed no detectable binding of siDNA up to NP concentrations of 10 mg/ml (Table 4-1). Interestingly, when primary and secondary groups were included on the same chain (A12), significant siDNA binding was observed (e.g., Table 4-1; compare VIII.14, $K = 236 \mu\text{g/ml}$, to VIII.43-VIII.46, $K > 10,000 \mu\text{g/ml}$). Moreover, binding was further enhanced by incorporation of an additional secondary amine group on the chain (A122) (Table 4-1; VIII.104, $K = 13.0 \mu\text{g/ml}$). We then reasoned that if we increased the polymer backbone length while maintaining the same percentage of amine along the backbone, we would detect improved binding because of the greater local positive charge density. Instead, we found that as backbone length increased, binding of siDNA decreased (Figure 4-5), perhaps indicating that the longer backbone polymers could not form the same structures when condensing the siDNA. As a standard for comparison for the binding affinities we were measuring, we performed the gel shift assay using the traditional polymers LPEI, BPEI, and poly-L-lysine (PLL) (Table 4-1). Our NPs with the highest measured binding affinities bound the siDNA nearly as tightly as these control polymers.

Polyethylene glycol (PEG) is commonly added to many of the polymers used for nucleic acid delivery because of its supposed benefits for charge shielding (Wolff and Rozema, 2008). We therefore included PEG on the majority of our NPs. Initially, either

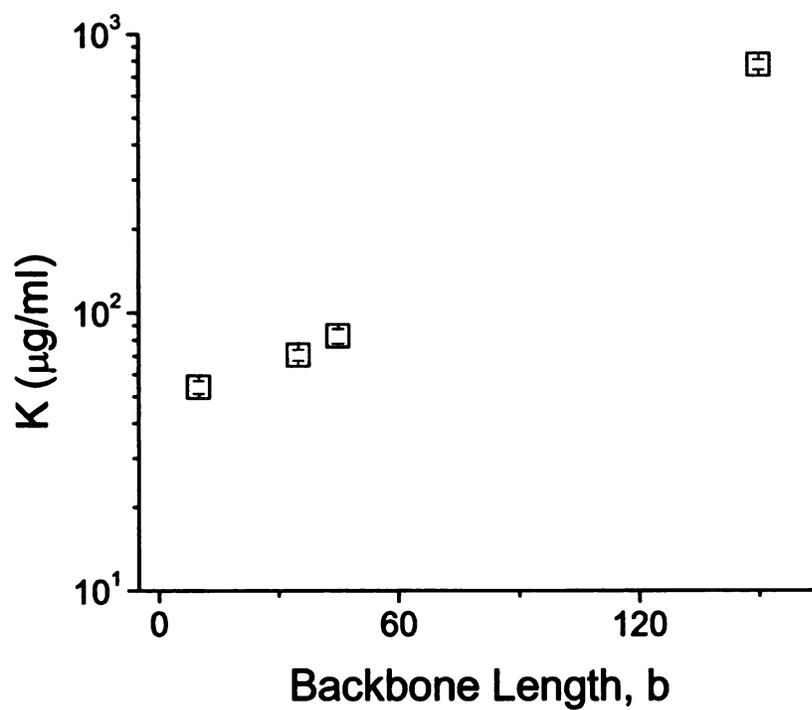


Figure 4-5 Nanoparticle binding affinity relative to polymer backbone length. NPs were synthesized with increasing backbone length and then tested for siDNA binding via gel shift assay. Binding affinity (K) was then plotted versus the backbone length. Better binding (smaller K values) correlated with shorter polymers.

8 or 115 repeat units of PEG were added directly to one end of the polymer backbone while keeping the compositions along the backbone constant. For a range of backbone compositions, PEG caused a reduction in binding affinity (Figure 4-6). This trend was also present when PEG groups were placed along the polymer backbone (Table 4.1; compare VIII.65, $K = 173 \mu\text{g/ml}$, to VIII.28, $K = 31.5 \mu\text{g/ml}$; and data not shown).

Having established that primary and secondary amines are required for siDNA binding, and that PEG decreased the overall binding affinity of the NPs, we next looked at how the remainder of the reactive alkyne sites could be filled to possibly improve binding or alter function. A series of NPs were synthesized that contained sites filled with a 0-, 4-, 10-, or 16-carbon (alkyl) chain such that all sites along the backbone were filled (i.e., the amine plus alkyl percent filled equals 100%; $y + z = 100$). Intriguingly, as the alkyl chain length increased, so too did the binding affinity (Figure 4-7). We hypothesize this result was due to the longer hydrophobic chains inducing vesicle-like formation around the siDNA, therefore making the NP better able to condense multiple siDNA molecules per complex. This was supported by NP size data measured with dynamic light scattering (DLS) where complex diameter was a linear function of alkyl length for the family of NPs with eight repeat units of PEG ($\text{PEG}_8\text{PG}_{45}(\text{A}_{1280}\text{C}_{20})$ where C was the type and length of the alkyl chain; Figure 4-8). Size measurements for the other NPs did not yield clear patterns (data not shown), perhaps because of the disruption of stable structures due to the longer PEG chains on those NPs. To get a second, independent confirmation of NP size and shape, we imaged the NPs alone or complexed with siDNA by transmission electron microscopy (TEM) and estimated the sizes from the images. As observed with DLS, the particles consistently ranged in size from ~15-150 nm, depending on the

particle chemistry (Figure 4-9). However, in all cases the NPs and complexes appeared to be spherical, as would be expected for vesicles.

Since our NPs were sufficiently interacting with siDNA and the complexes were on the nanometer scale making cell uptake feasible, we attempted to deliver siRNAs targeting the Enhanced Green Fluorescent Protein (EGFP) with our NPs into human lung cancer cells (H1299-EGFP) stably expressing the EGFP reporter plasmid. The siRNA was fluorescently tagged (red) so we could monitor delivery and localization. Although the siRNA was observed by fluorescence microscopy to enter the cells (Figure 4-2A), and the cells appeared healthy (Figure 4-2B), no silencing of the target EGFP was detected at any of the siRNA or NP concentrations tested (Figure 4-2C). At higher NP concentrations, the NP/siRNA complexes became toxic (Figure 4-10), with only a few chemistries leading to higher levels of toxicity than others (data not shown).

Our novel poly(propargyl glycolide) (PPGL) polymer backbone requires primary and secondary amine containing groups added to the backbone to facilitate siRNA binding. Indeed, when sufficient positive charge is present, the NPs can bind siRNA (or the analogous siDNA) at a level approaching what is observed with polymers currently in use in the literature. The complexes that form are on the nanometer scale where the size appears to depend on the alkyl chain length. More importantly, NP/siRNA complexes are readily taken up by the human cells tested here, but unfortunately, we have not yet observed gene silencing. One possibility is that our NPs are binding siRNA too strongly and the complexes are in such a conformation that release of the siRNA, either by charge neutralization or degradation of the polymer, does not occur on the time scale of the experiments. Further work will be necessary to determine why silencing is not observed

so that the polymers can be modified to overcome this problem. Overall, the results shown here will be useful in designing polymeric drug delivery systems for any nucleic acid cargo.

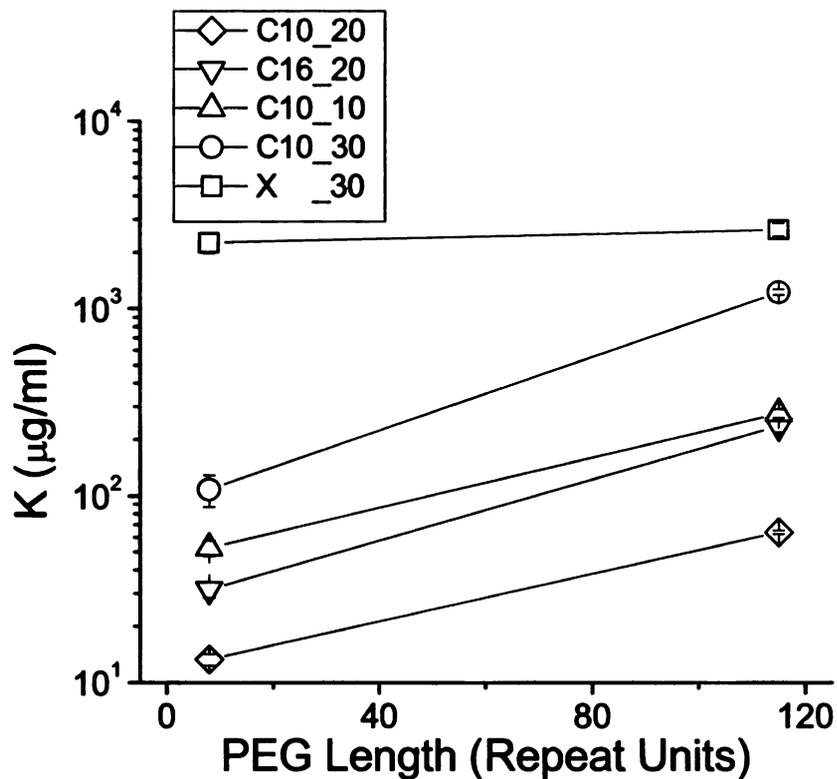


Figure 4-6 Nanoparticle binding affinity relative to PEG length.

NPs ($\text{PEG}_a\text{PG}_{45}(\text{A}_{12}_z\text{C}_y)$) containing the indicated alkyl groups (X, C10, or C16; $y = 10, 20, \text{ or } 30\%$) and amine group (A_{12} ; $z = 100\% - y$) were synthesized with PEG tails either 8 or 115 repeat units in length ($a = 8 \text{ or } 115$). Binding of siDNA was tested via gel shift assay and the affinity constant (K) was plotted versus the backbone length. Nomenclature is as described in Table 4-1.

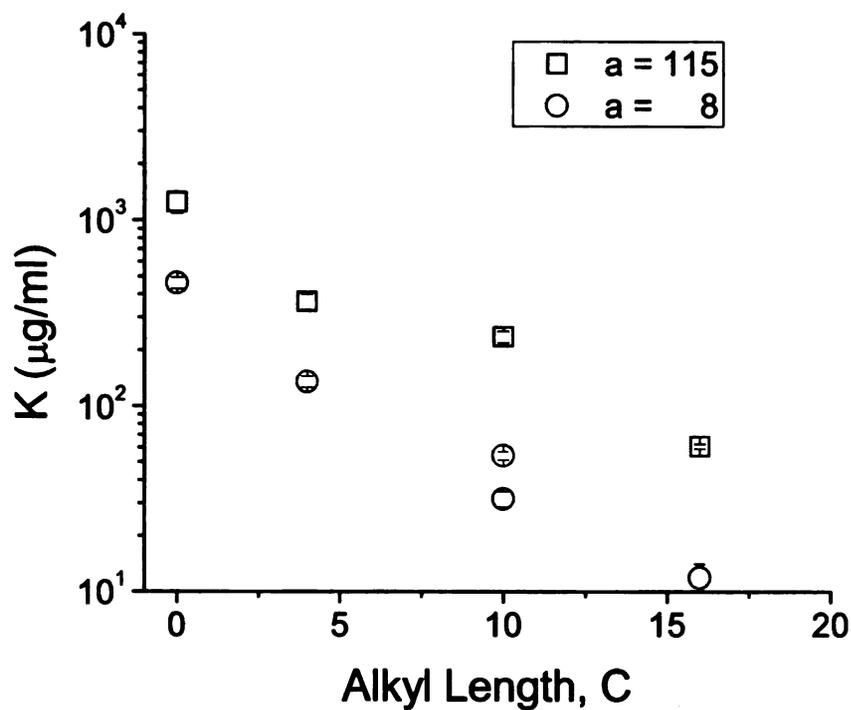


Figure 4-7 Nanoparticle binding affinity relative to alkyl chain length.

NPs ($\text{PEG}_a\text{PG}_{45}(\text{A}_{12}\text{zC}_y)$) were synthesized with amine group (A_{12} , $z = 80\%$) and alkyl chains of varying length ($C = 0, 4, 10, \text{ or } 16$, $y = 20\%$) attached to the polymer backbone. PEG tail length was varied ($a = 8$ or 115) as in Figure 4-6. Binding affinity (K) for siDNA was determined by gel shift assay and plotted versus the alkyl chain length (C).

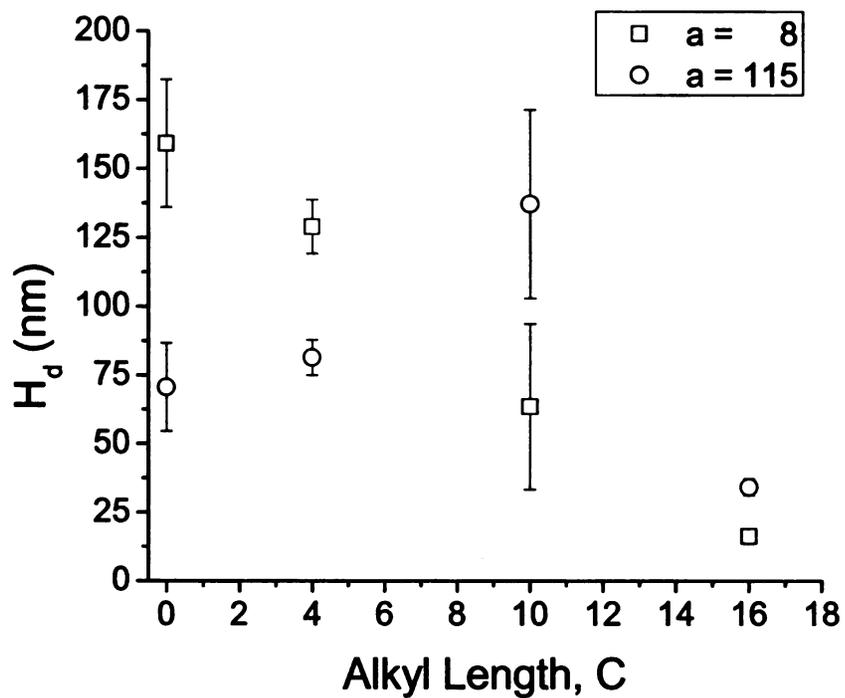
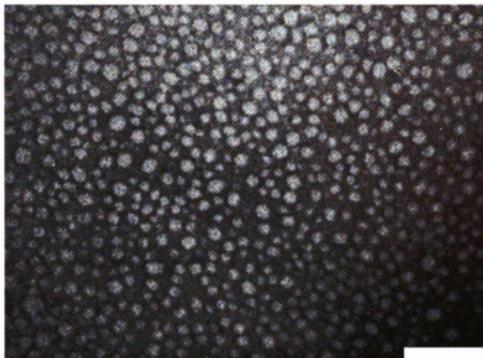


Figure 4-8 Nanoparticle size determined by Dynamic Light Scattering (DLS).

The average hydrodynamic diameter (H_d) of the NP/siDNA complexes ($\text{PEG}_a\text{PG}_{45}(\text{A}12_{80}\text{C}_{20})$) formed as in the gel shift assay were measured by DLS for NPs of varying alkyl lengths ($C = 0, 4, 10, \text{ or } 16$).

A



B

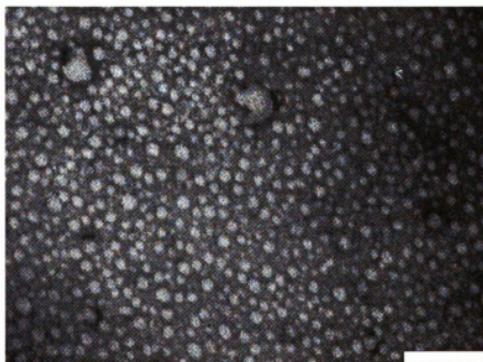


Figure 4-9 Nanoparticle size and shape determined by Transmission Electron Microscopy (TEM).

The NP/siDNA complexes were formed as in the gel shift assay and then imaged by TEM. Sizes of ~20-25 nm were estimated from the images for both (A) VIII.54 (PEG₈PG₄₅(A12₈₀X₂₀) and (B) VIII.71 (PEG₁₁₅PG₄₅(A12₈₀X₂₀). Images are at 200,000x magnification. Scale bars are 100 nm.

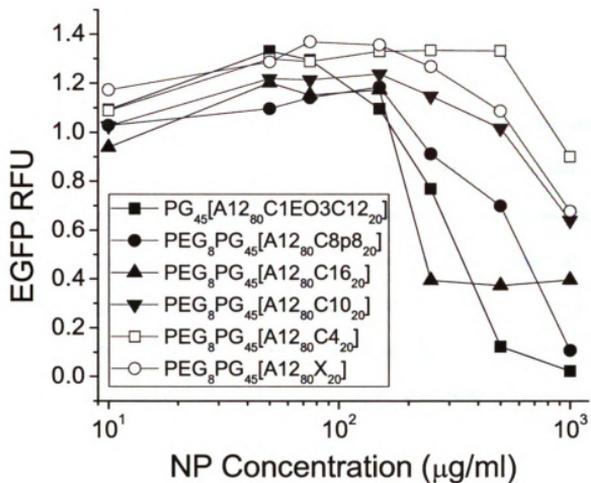


Figure 4-10 Nanoparticle toxicity.

The NPs were administered to human H1299 cells stably expressing EGFP for 24 hrs and then cell density was measured by quantification of EGFP levels. The NPs correspond to VIII.86 (■), VIII.68 (●), VIII.59 (▲), VIII.28 (▼), VIII.84 (□), and VIII.54 (○) as described in Table 4-1. RFU: Relative fluorescence units.

CHAPTER 5 CONCLUSIONS AND FUTURE WORK

5.1 Conclusions

The work presented here was motivated by the potential for siRNAs to be developed as therapeutics. Although much is already known about the role of siRNAs in RNAi, a better understanding of how siRNAs induce the pathway is needed before they can be of clinical use. To that end, it was shown that the secondary structure of the mRNA site targeted by siRNA has a significant effect on silencing activity. Sites that tended to lack structure were favorable targets, especially when the lack of structure was at either end of the target site. Additionally, siRNA guide strands predicted to have less structure were similarly preferred. Importantly, mRNA target site and siRNA guide strand secondary structures could be readily predicted by a free energy minimization algorithm. Therefore, this structure information can be incorporated into current siRNA selection algorithms with little effort and should help to improve the rate at which active siRNAs are identified.

The manner in which the guide strand of the siRNA is selected for loading onto Ago2 is also a critical step that can affect siRNA activity. However, in humans it has been unknown how that is accomplished. It was demonstrated here for the first time that human TAR RNA binding protein (TRBP), a component of RISC, can sense siRNA asymmetry, based most likely on both sequence and the overall and relative thermodynamic stability of the duplex. Surprisingly, TRBP interacted with ssRNAs in a fashion consistent with duplex recognition, providing further support for the role of TRBP in guide strand selection as the asymmetry sensor. Furthermore, it was shown that

TRBP interactions with siRNAs could be altered through the introduction of sequence mismatches, DNA substitutions, or 5'-O-methyl groups. Accordingly, manipulation of the siRNA by either sequence or chemical means offers a viable route for engineering improved siRNA function.

The creation of highly active siRNAs is paramount for eventual therapeutic applications and the results presented here regarding mRNA target structure effects and asymmetry sensing by TRBP offer two means for improving that process. However, an equally critical step is to achieve efficient delivery of the siRNA molecules to the cells of interest. A novel polymer system generated by Gregory Baker's laboratory via "click" chemistry was presented here and was used to study a range of variables contributing to siRNA binding, a key first step necessary for eventually achieving cellular uptake and *in vivo* silencing; such a systematic study has not been previously reported. It was found that high affinity binding of the nanoparticles with siRNA required a combination of primary and secondary amine groups on the same side chains off the polymer backbone, and that binding could be enhanced by simultaneously including alkyl side chains. Other changes in the polymer also altered the affinity, such as adding increasingly long PEG chains, either as a side chain or end-tail. Ultimately, uptake of siRNA into cells grown in culture could be achieved by the nanoparticles with minimal toxicity. Unfortunately, no target silencing was detected, indicating that further efforts will be required to optimize this new platform for function of the siRNA.

5.2 Future Work

The work presented here by no means solves all the problems related to the therapeutic application of siRNAs. Rather, many questions remain to be answered.

Below, I outline several topics that I feel are a direct extension of results included in this dissertation that if investigated, should help to answer some of those questions.

5.2.1 siRNA selection algorithm

Currently, the best siRNA selection algorithms incorporate numerous parameters that are both empirical and mechanistically relevant (insofar as the RNAi mechanism is presently understood). They tend to rely most heavily on rules that support incorporation of the appropriate guide strand, based on siRNA asymmetry, and account for mRNA target structure effects (Lu and Mathews, 2007; Shao et al., 2007). The results from Ch. 2 and 3 of this dissertation shed some light on both of these aspects of the RNAi pathway and therefore an attempt should be made to include them into an siRNA selection algorithm.

Our work here with TRBP suggests that siRNA asymmetry sensing is a real phenomenon that can be detected by RISC. However, it is unclear how siRNA asymmetry is sensed by TRBP and, more importantly, how best to predict the effect. Our results, and those of (Lu and Mathews, 2007), indicate that a single base-pair or nearest-neighbor energy parameter might be representative of asymmetry. Accordingly, more studies are warranted to determine which sequence or energy contributions most accurately correlate with activity so that they can be incorporated into the selection algorithm. Furthermore, as additional information is generated using chemical- and sequence-modified siRNAs with recombinant proteins (discussed below), it too can be included in the algorithm.

Downstream of siRNA asymmetry, RNA secondary structure effects become significant, as we and others have found (Ameres et al., 2007; Gredell et al., 2008; Lu

and Mathews, 2007; Shao et al., 2007). Although we did not do so, it would be interesting to see how our structure analyses would possibly differ if siRNA asymmetry were first taken into account. That is, we should repeat our guide strand and mRNA target site structure analyses after first classifying an siRNA as either asymmetric or symmetric (by whatever metric is found to be most significant).

It may also be desirable to repeat our RNA structure calculations using a partition function approach to determine the most likely structures instead of the minimum free energy structure, which may be present with limited frequency within a cell. While repeating the analysis, it is recommended that shorter mRNA sequences be used, perhaps ~100 nt up- and down-stream of the target site, to reduce the computational time required for the predictions, but also because it will likely give more realistic predictions (long-distance interactions are less likely).

A potential limitation of our approach used here is that the structure classification scheme is qualitative in nature, depending on classifications that are not at present quantifiable. Therefore, adapting our structure analysis technique to make the structure classifications quantitative may be worthwhile and may allow it to be easily coupled with the asymmetry calculation.

Lastly, there is the definite possibility that siRNA asymmetry and the guide strand and mRNA target site structure calculations will be insufficient alone to predict siRNA activity. As such, other variables frequently implemented in the literature should be considered for inclusion in this algorithm (e.g., GC content, internal stability, etc). The combination of these parameters could result in a viable algorithm founded primarily upon mechanistically relevant effects that would be of use for siRNA selection.

5.2.2 Further TRBP characterization

We have shown that TRBP alone is capable of sensing siRNA asymmetry and that binding appears to require at least some double-stranded, A-form RNA. But where exactly does TRBP bind an siRNA and how does it do so? The location of binding could be determined by repeating the UV-crosslinking experiments with siRNAs containing the 4-thiouracil group in different positions along the duplex. I would anticipate that TRBP most significantly interacts with the ends of the siRNA, with minimal interactions near the Ago2 cleavage site in the center of the duplex and consistent with what (Tomari et al., 2004) observed for R2D2. However, based on a model proposed by (Erard et al., 1998), I speculate that TRBP will hydrogen bond between lysine or arginine residues in TRBP and guanine nucleotides in the siRNA. Therefore, binding/crosslinking may depend on stretches of G:C base-pairs or the positioning of G:C pairs relative to each other.

It is also possible that TRBP nonspecifically recognizes the two nucleotide overhangs on either 3'-end of the siRNAs or the adjacent 5'-terminal nucleotide. Recognition of overhangs has previously been documented for the PAZ domain of Dicer (Ma et al., 2004; Ma et al., 2005), and it would not be surprising if TRBP did so as well. Our results here with the 4-thiouracil crosslinking agent located at position 20 (the first nucleotide in the overhang) suggest that TRBP is at least in close proximity to the overhangs and determining if the overhangs are necessary would easily be tested within the framework of our current experimental design by using siRNAs lacking overhangs.

Given the relative ease with which our binding/crosslinking experiments can be performed, a more systematic study could be undertaken to test additional siRNA chemical modifications (as outlined in Ch. 1) to identify their potential uses for dictating

siRNA asymmetry. Also of interest would be to see if TRBP can asymmetrically bind/crosslink to miRNAs to determine the role that TRBP plays in that related but albeit slightly different pathway. Collectively, this work should provide insight into how TRBP recognizes asymmetry and possibly highlight its broader role as a component of RISC.

5.2.3 Characterization of additional RNAi-related proteins

Human RISC is known to contain Dicer, TRBP, and Ago2 at a minimum (MacRae et al., 2008). However, several other proteins, including but not necessarily limited to, PACT, PKR, and RHA, may also interact directly with the complex and participate in the RNAi pathway. Many of the proteins contain dsRBDs that should mediate interactions with siRNAs, while others form protein-protein interactions directly through their respective C-terminal domains (the 'Medipal' domain) (Laraki et al., 2008). Although much is known about how these individual proteins function, their collective roles when combined into a multimeric protein complex during RNAi have not been studied.

The following experiments that I outline would require that some or all of these proteins be recombinantly expressed and purified. This task may not be as daunting as it initially appears. Our laboratory already has expressed and purified TRBP and PKR (from plasmids obtained from (Daher et al., 2001) and (Bevilacqua and Cech, 1996), respectively). PACT and RHA are similar to TRBP and PKR (they all contain multiple dsRBDs), which suggests that they could be relatively easily expressed and purified as either His- or MBP-tagged proteins, depending on their solubility. Instead, the most difficult tasks would likely be the purification of Dicer (it needs to be expressed in insect cells to ensure appropriate levels of post-translational modification) and Ago2 (it would

probably contain contaminating ssRNAs). However, both have previously been purified by (MacRae et al., 2008), suggesting that it is possible to obtain purified Dicer and Ago2.

The first experiments would be to determine if each of the proteins alone can sense siRNA asymmetry, as TRBP does. If the work proposed in Ch. 5.2.2 is carried out, this could logically be extended for the other proteins, as well, to clarify where and how the proteins recognize siRNA (or ssRNA). Based on the similarity of PACT, the activator of PKR, to TRBP (~42% identical) (Haase et al., 2005), there is a strong possibility that it too can sense siRNA asymmetry. This result would be particularly interesting if PACT provides an alternative means for selecting the siRNA guide strand. Also, since both TRBP and PACT have been shown to interact directly with Dicer, and in *Drosophila*, Dicer-2 appears to recognize the end opposite to R2D2 (Tomari et al., 2004), it would be straightforward to verify that a similar effect occurs in humans.

The second set of experiments would be designed to determine the temporal sequence of initial RLC formation, followed by final active RISC loading. It should be possible to determine which of the trio of RISC proteins (Dicer/TRBP/Ago2) first binds the siRNA, and where, and finally, when the guide strand gets loaded onto Ago2. The remaining proteins (PACT, RHA, or PKR) could then be added at various points to study their effects on the process. For example, RHA, as an RNA helicase, may enhance guide strand loading onto Ago2 by unwinding the siRNA duplex.

Taken together, these experiments would effectively reconstitute RISC and make it possible to clarify how the siRNA guide strand is initially loaded and subsequently used to dictate RISC targeting to mRNAs.

5.2.4 Further nanoparticle investigations

Just as it is important to develop a better understanding of how RISC forms and guides mRNA target cleavage, considerably more work is also needed if the polymeric nanoparticles we have developed are to be of value for delivering siRNAs *in vitro*. Our initial results show that the NP/siRNA complexes are taken up by the cells but do not cause appreciable reduction of the target protein levels. Therefore, additional studies are required to establish methods for improving silencing when initiated via NP-delivered siRNAs.

One possibility for achieving desirable levels of silencing is to characterize how the NP/siRNAs enter the cells so that cytoplasmic localization can be maximized. Polymers have been shown to enter cells through the endosomal pathway where their cargo can be trapped and prevented access to the cytosol (Douglas et al., 2008; Yezhelyev et al., 2008). Studies to track uptake typically utilize inhibitors of endosomes/endosomal acidification so that fluorescently tagged molecules can be monitored by confocal microscopy. These experiments could easily be repeated with our fluorescently tagged NPs or siRNAs. For NP/siRNA complexes that are detained in the endosomes, the polymers can be modified with pH sensitive moieties that can facilitate endosomal escape, such as was engineered by (Rozema et al., 2003; Rozema et al., 2007; Wakefield et al., 2005; Xiong et al., 2007), without causing excessive toxicity. After determining if our NPs are localized in the endosomes, we could begin to modify the polymer with these kinds of acid-labile groups to potentially enhance their release into the cytoplasm.

It is also possible that our NPs already enter the cytoplasm, and as a result of the improved siRNA binding we achieved through modifications to the polymer, we may have simultaneously made it more difficult for the siRNAs to be released. In that case, a systematic approach to investigating NP degradation/siRNA release is appropriate. Theoretically, these polymers are degradable into lactic acid derivatives (Gregory Baker, personal communication). But we do not know on what time scale this might occur. A useful study would be one that quantifies polymer degradation over time in various solutions, such as water, buffer, and possibly cell extract. It may be possible to similarly assess siRNA release from these complexes by performing gel shift assays and then monitoring for free siRNA (siDNA) over time.

As an alternative method for accelerating siRNA release, the siRNA could be attached directly to the polymer through disulfide bonds that would be subsequently reduced in the cytoplasm. This method has also been pursued by (Rozema et al., 2007) and is worth considering. Of course, if successful delivery and silencing can be achieved with any one of the related NPs derived in this study, a nearly endless list of experiments can be envisioned, perhaps culminating with an effective *in vivo* delivery study and eventually a clinical trial.

APPENDICES

Materials and Methods for Ch. 2

RNA secondary structure prediction and siRNA selection

The predicted secondary structure of the EGFP mRNA coding sequence was obtained using default settings on the mfold web server version 3.2 (<http://bioinfo.rpi.edu/applications/mfold>) or default settings on UNAFold version 3.4, an updated algorithm replacing mfold (Markham and Zuker, 2005; Zuker, 2003). siRNAs targeting the EGFP transcript were identified using Ambion Inc.'s siRNA Target Finder (http://www.ambion.com/techlib/misc/siRNA_finder.html). A subset of these siRNAs was further selected according to the structure of the mRNA target as predicted for the minimum free energy (MFE) structure (Figures 2-1 and 2-2). Chemically synthesized siRNAs with 3'-UU overhangs were purchased from Dharmacon. siRNA sequences and GC content are available in Table 2-1, and the predicted EGFP mRNA structures targeted by the siRNAs are shown in Figure 2-2. All siRNAs are referred to relative to the target region, where the 5'-end is position 1, which corresponds to position 19 of the antisense siRNA strand.

Cell culture and transfection

Human cervical carcinoma (HeLa) and human hepatocellular carcinoma (HepG2) cells were grown in Dulbecco's modified Eagle medium (DMEM; Invitrogen) supplemented with 10% (v/v) fetal bovine serum (Biomed), 100 U/ml penicillin and 100 µg/ml streptomycin (Invitrogen) at 37°C in a 5% CO₂ humidified incubator. For transfection, HeLa cells were plated in 6-well plates at 200,000 cells/well and HepG2

cells were plated in 12-well plates at 400,000 cells/well 24 h before use in medium containing serum but lacking antibiotic. EGFP plasmid (4.0 μ g for HeLa, 1.6 μ g for HepG2, d2EGFP (EGFP variant with an \sim 2 hr half-life), Clontech) and 30 nM siRNA were transfected using Lipofectamine 2000 (3.5 and 6.0 μ l, respectively; Invitrogen) in OptiMEM (Invitrogen) according to the manufacturer's protocol. The transfection solution was replaced after four hours with complete culture medium.

Transfection efficiency was monitored by fluorescence microscopy 24 hrs post-transfection using fluorescently tagged siRNAs (targeting position 126 of the EGFP transcript or a non-targeting sequence (Block-iT Fluorescent Oligo; Invitrogen)). In all cases, transfection efficiency was $> 85\%$. Consistent cell viability, independent of siRNA and/or EGFP plasmid transfection, was confirmed by both microscopy and CellTiter-Blue Cell Viability Assay (Promega) performed according to the manufacturer's instructions. Efficiency of EGFP plasmid transfection was similarly monitored by microscopy. EGFP expression and cell viability were consistent throughout the experiments.

EGFP protein expression quantification

EGFP expression was quantified from live cells 24 hrs post-transfection. Culture medium was aspirated, cells were washed twice with phosphate buffered saline (PBS), and then a volume of PBS equal to the culture medium volume was added. Finally, fluorescence levels were measured using a SPECTRAmax GEMINI EM plate reader (Molecular Devices) with excitation at 480 nm, emission at 525 nm, and cutoff filter at 515 nm, as recommended by the manufacturer. Relative fluorescence units (RFU) were scaled to EGFP transfected cells with no siRNA (100%) and mock transfected cells with

no siRNA or plasmid (0%). Fluorescence levels from mock transfected cells differed by less than 2% compared to wells containing only PBS. Experiments were repeated at least eight times ($n \geq 8$).

Data and analysis

RNA secondary structure predictions were performed on the genes targeted by siRNAs in the dataset compiled by (Shabalina et al., 2006) that contained 653 siRNAs. Prior to the analysis, we removed siRNAs targeting four non-human genes and deleted five genes that contained over 6,000 nts, as this was the maximum length that could be analyzed by mfold. Two other genes were removed due to difficulty in identifying the appropriate target mRNA sequences. The final dataset consisted of 548 siRNAs (533 from (Shabalina et al., 2006) and 15 from this work) targeting 42 different genes. As such, we expect that the pool of sequences used in our analyses is still sufficiently large and targets enough different mRNAs to be representative of the entire siRNA target landscape.

Structures targeted by siRNAs were identified from the MFE prediction of the full-length mRNA sequence, as listed by the National Center for Biotechnology Information (NCBI) on 12 April 2007, and analyzed as described in Ch. 2.3. The MFE secondary structure of each siRNA guide strand was also determined using default mfold settings and analyzed as described. Due to the variability in siRNA overhang composition utilized throughout the literature (i.e., 3'-UU vs. 3'-dTdT vs. matching the target site), only the effects of the 19 nucleotide core on target structure were considered. All other calculations including unpaired Student's t-test to compare two independent groups were performed using Microsoft Excel.

Materials and Methods for Ch. 3

Protein expression and purification

Plasmids encoding TRBP as a fusion product with maltose binding protein (MBP-TRBP) and MBP alone were kindly provided by Dr. Anne Gatignol and expressed essentially as described (Laraki et al., 2008). Briefly, plasmids were transformed into Rosetta2(DE3) competent cells (Novagen). An overnight culture from a single colony was diluted 1:50 (v/v) and grown for 3-4 hrs at 37°C, or until $A_{600} = 0.6-1.0$, and then induced with 0.3 mM IPTG. After 2 hrs of expression at 37°C, cells were pelleted (4000 rpm for 10 min), resuspended in Column Buffer (20 mM Tris-HCl, 200 mM NaCl, 1 mM EDTA), lysed via sonication, and then clarified by high speed centrifugation (15,000 rpm for 25 min). The supernatant was purified with an AKTA FPLC (GE Healthcare). TRBP-MBP and MBP were eluted from a MBPTrap Column (GE Healthcare) with 5 column volumes of Elution Buffer (20 mM Tris-HCl, 200 mM NaCl, 1 mM EDTA, 10 mM Maltose) and stored in Elution Buffer at -80°C until use.

Protein concentration was measured by Bradford assay (BioRad) or Bicinchoninic acid (BCA) assay (Pierce; Thermo Scientific). Products were assessed by separating approximately equal amounts of protein per lane on denaturing linear 4-20% Tris-HCl gels. Total protein was visualized by staining with Gel Code Blue (Pierce; Thermo Scientific) and specific proteins were detected via western blot. For western blot, the separated proteins were transferred to nitrocellulose membranes and then incubated overnight at 4°C with the appropriate antibodies (MBP (New England Biolabs), TRBP (Abnova)). Blots were washed with TBS-Tween, incubated for an additional 1 hr with HRP-linked secondary antibody (Pierce; Thermo Scientific), and developed using the

SuperSignal West Femto Maximum Sensitivity Substrate kit (Pierce; Thermo Scientific).

Images were obtained on a BioRad ChemiDoc XRS imager using Quantity One software.

Nucleic acids

DNA oligonucleotides were purchased from Integrated DNA Technologies. Chemically synthesized RNAs were purchased from Dharmacon as either ready-to-use siRNAs or ssRNAs. The lists of sequences used are provided in Table 6-1 to 6-5. HIV-1 TAR RNA was prepared by *in vitro* transcription from a plasmid kindly provided by Dr. K-T Jeang. Briefly, 1 μ g of linearized plasmid DNA was transcribed with the Megashortscript Kit (Ambion) following the manufacturer's instructions. Transcripts were resolved in 8M urea-10% polyacrylamide gels, visualized by UV shadowing, and eluted from gel slices by crushing and soaking in TE buffer (10 mM Tris, 1 mM EDTA) for 10 min at 75°C. RNA was then concentrated by ethanol precipitation, washed with 70% ethanol, and quantitated spectrophotometrically (NanoDrop). In order to remove the 5'-triphosphate, RNAs were treated with calf intestinal phosphatase (New England Biolabs). Treated RNAs were purified first by phenol/chloroform extraction, then ethanol precipitated, gel purified (as above), and concentrated. The concentration of the product containing 5'-OH was again measured prior to end-labeling with [γ -³²P]-ATP (see below).

Table 6-1 Nucleic acids used in Ch. 3 (listed 5' to 3').

Strand Name	Sequence
pp-luc-ss	CGUACGCGGAAUACUUCGAUA
pp-luc-ss-4SU(20)	CGUACGCGGAAUACUUCGAUA
pp-luc-as	UCGAAGUAUUCGCGUACGUG
pp-luc-as-4SU(20)	UCGAAGUAUUCGCGUACGUG
pp-luc-ss-A(1)	<u>A</u> GUACGCGGAAUACUUCGAUA
pp-luc-ss-C(19)	CGUACGCGGAAUACUUCG <u>C</u> UA
sod1-as	GUCACAUUGCCCAAGUCUCUU
sod1-as-4SU(20)	GUCACAUUGCCCAAGUCUCUU
sod1-ss	GAGACUUGGGCAAUGUGACUU
sod1-ss-4SU(20)	GAGACUUGGGCAAUGUGACUU
sod1-as-U(1)	<u>U</u> AGACUUGGGCAAUGUGACUU
sod1-as-A(19)	GAGACUUGGGCAAUGUGA <u>A</u> UU
EGFP-as	UGCGCUCCUGGACGUAGCCUU
EGFP-as-4SU(20)	UGCGCUCCUGGACGUAGCCUU
EGFP-ss	GGCUACGUCCAGGAGCGCAUU
EGFP-ss-4SU(20)	GGCUACGUCCAGGAGCGCAUU
EGFP-ss-U(1)	<u>U</u> GCUACGUCCAGGAGCGCAUU
EGFP-ss-C(19)	GGCUACGUCCAGGAGCGC <u>C</u> UU
EGFP-as-5'OCH3	mTGCGCUCCUGGACGUAGCCUU
EGFP-ss-5'OCH3	mTGCUACGUCCAGGAGCGCAUU

U denotes 4-thiouracil; mT denotes 5'-O-methyl thymidine; underline denotes nucleotide that is mismatched to its complementary sequence.

Table 6-2 Nucleic acid duplexes used in Ch. 3 (top strand listed 5' to 3', bottom strand listed 3' to 5').

Strand Name	Sequence
RNA _{ss}	p* - GGCUACGUCCAGGAGCGCAUU
RNA _{as}	UCCGAUGCAGGUCCUCGCGU - p*
DNA _{ss}	p* - GGCTACGTCCAGGAGCGCATT
DNA _{as}	TTCCGATGCAGGTCCTCGCGT - p*
RNA _{ss}	p* - GGCUACGUCCAGGAGCGCAUU
DNA _{as}	TTCCGATGCAGGTCCTCGCGT - p*
DNA _{ss}	p* - GGCTACGTCCAGGAGCGCATT
RNA _{as}	UCCGAUGCAGGUCCUCGCGU - p*

* denotes 5'-phosphorylation with γ -³²P.

Table 6-3 Nucleic acid duplexes targeting pp-luciferase used in Ch. 3 (top strand listed 5' to 3', bottom strand listed 3' to 5').

Duplex Name	Strand Name	Sequence
a	pp-luc-as	p* -UCGAAGUAUCCGCGUACGUG
	pp-luc-ss	AUAGCUUCAUAAGGCGCAUGC -p*
b	pp-luc-as-4SU(20)	p* -UCGAAGUAUCCGCGUACGUG
	pp-luc-ss	AUAGCUUCAUAAGGCGCAUGC -p
c	pp-luc-as-4SU(20)	p -UCGAAGUAUCCGCGUACGUG
	pp-luc-ss	AUAGCUUCAUAAGGCGCAUGC -p*
d	pp-luc-as	p -UCGAAGUAUCCGCGUACGUG
	pp-luc-ss-4SU(20)	AUAGCUUCAUAAGGCGCAUGC -p*
e	pp-luc-as	p* -UCGAAGUAUCCGCGUACGUG
	pp-luc-ss-4SU(20)	AUAGCUUCAUAAGGCGCAUGC -p
f	pp-luc-as-4SU(20)	p* -UCGAAGUAUCCGCGUACGUG
	pp-luc-ss-4SU(20)	AUAGCUUCAUAAGGCGCAUGC -p*
g	pp-luc-as-4SU(20)	p* -UCGAAGUAUCCGCGUACGUG
	pp-luc-ss-A(1)	AUAGCUUCAUAAGGCGCAUGA -P
h	pp-luc-as-4SU(20)	p* -UCGAAGUAUCCGCGUACGUG
	pp-luc-ss-C(19)	AUCGCUUCAUAAGGCGCAUGC -P
i	pp-luc-as-4SU(20)	p* -UCGAAGUAUCCGCGUACGUG
	<i>pp-luc-ss</i>	<i>ATAGCTTCATAAGGCGCATGC -p</i>
j	<i>pp-luc-as</i>	<i>p -TCGAAGTATTCGCGTACGTG</i>
	pp-luc-ss-4SU(20)	AUAGCUUCAUAAGGCGCAUGC -p*

U denotes 4-thiouracil; underline denotes nucleotide that is mismatched to its complementary sequence; italics denote DNA strands; p denotes 5'-phosphorylation; * denotes 5'-phosphorylation with γ -³²P.

Table 6-4 Nucleic acid duplexes targeting *sod1* used in Ch. 3 (top strand listed 5' to 3', bottom strand listed 3' to 5').

Duplex Name	Strand Name	Sequence
a	<i>sod1</i> -ss	p* -GAGACUUGGGCAAUGUGACUU
	<i>sod1</i> -as	UUCUCUGAACCCGUUACACUG-P*
b	<i>sod1</i> -ss-4SU(20)	p* -GAGACUUGGGCAAUGUGAC <u>UU</u>
	<i>sod1</i> -as	UUCUCUGAACCCGUUACACUG-P
c	<i>sod1</i> -ss-4SU(20)	p -GAGACUUGGGCAAUGUGAC <u>UU</u>
	<i>sod1</i> -as	UUCUCUGAACCCGUUACACUG-P*
d	<i>sod1</i> -ss	p -GAGACUUGGGCAAUGUGACUU
	<i>sod1</i> -as-4SU(20)	<u>UUCUCUGAACCCGUUACACUG</u> -P*
e	<i>sod1</i> -ss	p* -GAGACUUGGGCAAUGUGACUU
	<i>sod1</i> -as-4SU(20)	<u>UUCUCUGAACCCGUUACACUG</u> -P*
f	<i>sod1</i> -ss-4SU(20)	p* -GAGACUUGGGCAAUGUGAC <u>UU</u>
	<i>sod1</i> -as-4SU(20)	<u>UUCUCUGAACCCGUUACACUG</u> -P*
g	<i>sod1</i> -ss-4SU(20)	p* -GAGACUUGGGCAAUGUGAC <u>UU</u>
	<i>sod1</i> -as-U(1)	UUCUCUGAACCCGUUACAC <u>UU</u> -P
h	<i>sod1</i> -ss-4SU(20)	p* -GAGACUUGGGCAAUGUGAC <u>UU</u>
	<i>sod1</i> -as-A(19)	UU <u>A</u> UCUGAACCCGUUACACUG-P
i	<i>sod1</i> -ss-4SU(20)	p* -GAGACUUGGGCAAUGUGAC <u>UU</u>
	<i>sod1</i> -as	UUCUCUGAACCCGUUACACUG-P
j	<i>sod1</i> -ss	p -GAGACUUGGGCAAUGUGACUU
	<i>sod1</i> -as-4SU(20)	<u>UUCUCUGAACCCGUUACACUG</u> -P*

U denotes 4-thiouracil; underline denotes nucleotide that is mismatched to its complementary sequence; p denotes 5'-phosphorylation; * denotes 5'-phosphorylation with γ -³²P.

Table 6-5 Nucleic acid duplexes targeting EGFP used in Ch. 3 (top strand listed 5' to 3', bottom strand listed 3' to 5').

Duplex Name	Strand Name	Sequence
a	EGFP-as	p* -UGCGCUCCUGGACGUAGCCUU
	EGFP-ss	UUACGCGAGGACCUGCAUCGG-p*
b	EGFP-as-4SU(20)	p* -UGCGCUCCUGGACGUAGCCUU
	EGFP-ss	UUACGCGAGGACCUGCAUCGG-p
c	EGFP-as-4SU(20)	p -UGCGCUCCUGGACGUAGCCUU
	EGFP-ss	UUACGCGAGGACCUGCAUCGG-p*
d	EGFP-as	p -UGCGCUCCUGGACGUAGCCUU
	EGFP-ss-4SU(20)	UUACGCGAGGACCUGCAUCGG-p*
e	EGFP-as	p* -UGCGCUCCUGGACGUAGCCUU
	EGFP-ss-4SU(20)	UUACGCGAGGACCUGCAUCGG-p
f	EGFP-as-4SU(20)	p* -UGCGCUCCUGGACGUAGCCUU
	EGFP-ss-4SU(20)	UUACGCGAGGACCUGCAUCGG-p*
g	EGFP-as-4SU(20)	p* -UGCGCUCCUGGACGUAGCCUU
	EGFP-ss-U(1)	UUACGCGAGGACCUGCAUCGU-p
h	EGFP-as-4SU(20)	p* -UGCGCUCCUGGACGUAGCCUU
	EGFP-ss-C(19)	UU <u>CC</u> GCGAGGACCUGCAUCGG-p
k	EGFP-as-4SU(20)	p* -UGCGCUCCUGGACGUAGCCUU
	EGFP-ss-5'OCH3	UUACGCGAGGACCUGCAUCGTm
l	EGFP-as-5'OCH3	mTGCGCUCCUGGACGUAGCCUU
	EGFP-ss-4SU(20)	UUACGCGAGGACCUGCAUCGG-p*

U denotes 4-thiouracil; mT denotes 5'-O-methyl thymidine; underline denotes nucleotide that is mismatched to its complementary sequence; italics denote DNA strands; p denotes 5'-phosphorylation; * denotes 5'-phosphorylation with γ -³²P.

siRNA duplex thermodynamic calculation

The method for calculating the relative thermodynamic asymmetry for siRNA duplexes was performed as described by (Hutvagner, 2005) using the most recent nearest-neighbor values available (Mathews et al., 1999). Basically, the free energy (ΔG) was calculated for each end of the duplex by summing the four terminal nearest-neighbor values (stacking energies), the value for the first nucleotide of the overhang (single base stacking energies), and a penalty of +0.5 kcal/mol for a terminal A:U pairing (miscellaneous energies). The relative asymmetry ($\Delta\Delta G$) was then calculated by taking the difference between the free energy of the AS 5'-end and the free energy of the SS 5'-end. For example, for the asymmetric pp-luc siRNA, the AS 5'-end $\Delta G = (-2.40) + (-2.40) + (-2.40) + (-0.90) + (-0.60) + (0.5) = -8.2$ kcal/mol, the SS 5'-end $\Delta G = (-2.40) + (-2.20) + (-1.30) + (-2.20) + (-0.6) + (0.0) = -8.7$ kcal/mol, yielding a $\Delta\Delta G = (-8.2) - (-8.7) = 0.5$ kcal/mol. Since $\Delta\Delta G > 0$, it indicates that the SS 5'-end is more stably hybridized than the AS 5'-end, thereby favoring incorporation of the AS into RISC.

Duplex preparation and end-labeling

For Figure 3-2, oligonucleotides were hybridized by mixing equal amounts of both top and bottom strands in 1x STE buffer (10 mM Tris, 100 mM NaCl, 1 mM EDTA) and then heated to 90°C for 3 min, followed by incubation at 37°C for 60 min. Products were verified on native TBE gels stained with SYBR Gold (Invitrogen). Duplexes for Figure 3-2 and ssRNAs (Figure 3-11) (3 pmol) were directly 5'-radiolabeled with 10 pmol of [γ -³²P]-ATP using T4 polynucleotide kinase (New England Biolabs); unincorporated label was removed using G-25 Sephadex columns (Roche Applied Science). TAR RNA (3 pmol) was similarly labeled following CIP treatment.

For the remaining figures in Ch. 3 where strands were alternately hot and cold labeled, individual single-strands (3 pmol) were labeled as described above (10 pmol of ATP) in 25 μ l reactions, where nonisotopic ATP was used for cold labeling. The 5'-O-methyl modified strands were not end-labeled. The two strands comprising the siRNA duplex were then mixed (50 μ l), 5.56 μ l of 10x STE was added for 1.0x STE final concentration, and samples were heated to 90°C for 3 min and then 37°C for 60 min. Unincorporated label was then removed from the annealed siRNA using the Sephadex columns. Final products were verified on native TBE gels.

Native gel shift assay

TRBP binding was determined by gel shift assays using radiolabeled oligonucleotides present at limiting concentrations relative to protein concentrations. Binding reactions (10 μ l) were prepared with $\sim 1-3 \times 10^4$ cpm of nucleic acid and 350 nM TRBP (unless noted otherwise) in Elution Buffer supplemented with 20 mM HEPES, 40 mM KCl, 1.5 mM MgCl₂, 0.1% Nonidet P40, and 10 U SUPERase-In (Ambion). Samples were incubated at room temperature for 30 min, mixed with 2 μ l of 5x Nucleic Acid Sample Loading Buffer (BioRad), and 10 μ l were loaded onto a pre-run native TBE gel. Mini-gels were prepared where the top half of the gel was 4% (37:1 acrylamide:bisacrylamide) and the bottom half was a gradient between 4-20%. Gels were run at 150 V for ~ 40 min, dried on filter paper under vacuum at 80°C for 50 min, exposed to a storage phosphor screen overnight ($\sim 12-16$ hrs), and then imaged on a Storm 860 imager (GE Healthcare). Relative signal intensities were quantified using Quantity One software and normalized within a single lane. Native gels for loading

control were similarly prepared but with the volume of TRBP replaced with an equal volume of Elution Buffer.

UV crosslinking, denaturing gel shift assay

TRBP crosslinking to nucleic acids was measured as in the native gel shift assay except that after the initial 30 min binding, the samples were subsequently exposed to 312nm UV light (Fisher Scientific Electrophoresis Systems Transilluminator) for 10 min while on an ice cold aluminum block (4°C). In order to minimize non-specific crosslinking, the samples were covered by a Petri dish to block shorter wavelengths. Samples were then mixed with 5 µl of 3x SDS Sample Buffer (New England Biolabs), boiled at 95°C for 3 min, collected by brief centrifugation, and 12 µl were loaded onto a pre-run non-linear denaturing 4-20% Tris-HCl gel. Gels were run, dried, imaged, and analyzed as described above. Statistical analyses are reported as *p*-values from unpaired, two-tailed Student's *t*-test performed using Microsoft Excel.

Materials and Methods for Ch. 4

Nucleic acids

Three types of nucleic acids were considered: deoxyribonucleic acids (DNA), plasmid DNA (pDNA), and short interfering ribonucleic acids (siRNA). DNA was chemically synthesized by Integrated DNA Technologies (IDT) as 21 nt long sequences (5'-CTGGGTGCTCAGGTAGTGGTT-3', sense strand; 5'-CCACTACCTGAGCACCCAGTT-3', antisense strand) that were complementary. When hybridized, the dsDNA product was 21 base-pairs long with TT overhangs on both 3'-ends, similar to the structure characteristic of siRNAs. When noted, the DNA was

modified with a 6-FAM on the 5'-end so that it would provide green fluorescence when excited by UV light. Plasmid DNA encoding the destabilized version of the Enhanced Green Fluorescent Protein (EGFP; d2EGFP, Clontech) such that the protein half-life was ~2hrs was prepared following standard protocols. siRNAs targeting the EGFP were chemically synthesized by Dharmacon, Inc. and the sequences have been published previously (Gredell et al., 2008). A non-targeting siRNA containing a FITC modification on both 5'-ends for green fluorescence imaging was purchased from Invitrogen (Block-iT Fluorescent Oligo).

Polymeric nanoparticle synthesis

Nanoparticles (NPs) were synthesized following the methods described in (Jiang et al., 2008) and in more detail elsewhere (Gregory Baker, personal communication).

Gel shift assay

An siDNA 21 nt in length and possessing 2 nt overhangs on either 3'-end (siDNA), analogous to an siRNA, was prepared by annealing complementary strands according to standard techniques described in Appendices – Materials and Methods for Ch. 3. To simplify imaging, both strands of the siDNA contained 5'-6-FAM modifications.

NP or control polymer binding of siDNA was measured using a gel shift assay. Briefly, 25 μ l reactions were prepared in water with 200 nM dsDNA and NP concentrations typically ranging from 0.10-7500 μ g/mL. Reactions were incubated at room temperature for 15 min before addition of 6.25 μ l of 5x Nucleic Acid Loading Buffer (BioRad). An aliquot of each sample (15 μ l) was loaded onto a 0.8% agarose gel (prepared in 45 mM Tris-Borate, 1 mM EDTA buffer) and electrophoresed at 50 V for

~35 min, or until the first blue front was ~2/3 down the gel. Gels were imaged and analyzed on a BioRad Chemidoc XRS machine. Band intensities for unbound siDNA from lanes with NP were normalized to lanes containing only siDNA to determine the fraction of unbound siDNA (fraction unbound = [intensity of unbound siDNA from NP treated lane]/[intensity of untreated siDNA]). The fraction of bound siDNA was calculated from the unbound fraction (fraction bound siDNA = 1 – fraction unbound). Binding constants (K) were determined for each NP by fitting the fraction of siDNA bound versus NP concentration to the Hill equation using Origin (Microcal Software, Inc.).

Dynamic light scattering and transmission electron microscopy

NP/siDNA complex sizes were measured using dynamic light scattering (DLS) and transmission electron microscopy (TEM) by first forming complexes as in the gel shift assay and then following standard methods to be described elsewhere that were specific to the particular technique (Gregory Baker, personal communication).

Application of nanoparticle/siRNA complexes to cell culture

The cellular response to NP or NP/siRNA complexes was evaluated by treating the NP as if it were a lipid-based transfection and then applying it to cells grown in culture according to standard transfection techniques (see Appendices – Materials and Methods for Ch. 2 or www.invitrogen.com for details).

Initial validation of rhodamine-tagged NP uptake was performed after incubating HeLa (human cervical carcinoma) cells with NP mixtures up to ~1.0 mg/ml for 24 hrs. Cells were washed to remove NPs remaining in solution and then imaged by fluorescence

microscopy to monitor the localization and intensity of rhodamine (red) signal (Figure 4-2A).

The NP chemistries that were capable of binding siDNA were tested for their effectiveness to facilitate siRNA uptake into H1299 cells (kindly provided by Prof. Claus Bus) stably expressing an EGFP plasmid related to the one described in Appendices – Materials and Methods for Ch. 2. A Dy547-modified siRNA synthesized by Dharmacon (30 nM) and known to effectively silence the target EGFP mRNA was mixed with ~1.0-5,000 $\mu\text{g/ml}$ NP as in the gel shift assay using filter sterilized ($< 0.22 \mu\text{m}$) solutions whenever possible. Total volumes were scaled according to the volume of medium in which the cells were cultured to achieve the desired final concentrations. NP/siRNA complexes were applied to the H1299 cells for 24 hrs and then cells were imaged by fluorescence microscopy to assess EGFP knockdown (loss of green signal) and siRNA localization (red signal from Dy547). To measure NP or NP/siRNA toxicity profiles, the EGFP signal from the H1299 cells was used as an indicator of total cell number. H1299 cells grown in 96-well plates were treated in triplicate with NP or NP complexed with a non-targeting siRNA and EGFP levels were quantified from live cells as in Appendices – Materials and Methods for Ch. 2. EGFP relative fluorescence units (RFU) were calculated by normalizing levels for treated samples to untreated control samples.

BIBLIOGRAPHY

- Akinc, A., Zumbuehl, A., Goldberg, M., Leshchiner, E.S., Busini, V., Hossain, N., Bacallado, S.A., Nguyen, D.N., Fuller, J., Alvarez, R., *et al.* (2008). A combinatorial library of lipid-like materials for delivery of RNAi therapeutics. *Nat Biotechnol* 26, 561-569.
- Allen, R.S., Miller, J.A., Chitty, J.A., Fist, A.J., Gerlach, W.L., and Larkin, P.J. (2008). Metabolic engineering of morphinan alkaloids by over-expression and RNAi suppression of salutaridinol 7-O-acetyltransferase in opium poppy. *Plant Biotechnol J* 6, 22-30.
- Allen, R.S., Millgate, A.G., Chitty, J.A., Thisleton, J., Miller, J.A., Fist, A.J., Gerlach, W.L., and Larkin, P.J. (2004). RNAi-mediated replacement of morphine with the nonnarcotic alkaloid reticuline in opium poppy. *Nat Biotechnol* 22, 1559-1566.
- Amarzguioui, M., and Prydz, H. (2004). An algorithm for selection of functional siRNA sequences. *Biochem Biophys Res Commun* 316, 1050-1058.
- Ameres, S.L., Martinez, J., and Schroeder, R. (2007). Molecular basis for target RNA recognition and cleavage by human RISC. *Cell* 130, 101-112.
- Bartlett, D.W., and Davis, M.E. (2007a). Effect of siRNA nuclease stability on the in vitro and in vivo kinetics of siRNA-mediated gene silencing. *Biotechnol Bioeng* 97, 909-921.
- Bartlett, D.W., and Davis, M.E. (2007b). Physicochemical and biological characterization of targeted, nucleic acid-containing nanoparticles. *Bioconjug Chem* 18, 456-468.
- Baulcombe, D. (2004). RNA silencing in plants. *Nature* 431, 356-363.
- Benkirane, M., Neuveut, C., Chun, R.F., Smith, S.M., Samuel, C.E., Gatignol, A., and Jeang, K.T. (1997). Oncogenic potential of TAR RNA binding protein TRBP and its regulatory interaction with RNA-dependent protein kinase PKR. *EMBO J* 16, 611-624.
- Berkhout, B., and Jeang, K.T. (1991). Detailed mutational analysis of TAR RNA: critical spacing between the bulge and loop recognition domains. *Nucleic Acids Res* 19, 6169-6176.
- Bernstein, E., Caudy, A.A., Hammond, S.M., and Hannon, G.J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* 409, 363-366.
- Bevilacqua, P.C., and Cech, T.R. (1996). Minor-groove recognition of double-stranded RNA by the double-stranded RNA-binding domain from the RNA-activated protein kinase PKR. *Biochem* 35, 9983-9994.

- Bevilacqua, P.C., George, C.X., Samuel, C.E., and Cech, T.R. (1998). Binding of the protein kinase PKR to RNAs with secondary structure defects: role of the tandem A-G mismatch and noncontiguous helices. *Biochem* 37, 6303-6316.
- Bohula, E.A., Salisbury, A.J., Sohail, M., Playford, M.P., Riedemann, J., Southern, E.M., and Macaulay, V.M. (2003). The Efficacy of Small Interfering RNAs Targeted to the Type 1 Insulin-like Growth Factor Receptor (IGF1R) Is Influenced by Secondary Structure in the IGF1R Transcript. *J Biol Chem* 278, 15991-15997.
- Bridge, A.J., Pebernard, S., Ducraux, A., Nicoulaz, A.-L., and Iggo, R. (2003). Induction of an interferon response by RNAi vectors in mammalian cells. *Nat Genet* 34, 263-264.
- Brown, K.M., Chu, C.Y., and Rana, T.M. (2005). Target accessibility dictates the potency of human RISC. *Nat Struct Mol Biol* 12, 469-470.
- Brus, C., Petersen, H., Aigner, A., Czubayko, F., and Kissel, T. (2004). Physicochemical and Biological Characterization of Polyethylenimine-graft-Poly(ethylene glycol) Block Copolymers as a Delivery System for Oligonucleotides and Ribozymes. *Bioconjug Chem* 15, 677-684.
- Carlson, C.B., Stephens, O.M., and Beal, P.A. (2003). Recognition of double-stranded RNA by proteins and small molecules. *Biopolymers* 70, 86-102.
- Caudy, A.A., and Hannon, G.J. (2004). Induction and biochemical purification of RNA-induced silencing complex from *Drosophila* S2 cells. *Methods Mol Biol* 265, 59-72.
- Caudy, A.A., Ketting, R.F., Hammond, S.M., Denli, A.M., Bathoorn, A.M., Tops, B.B., Silva, J.M., Myers, M.M., Hannon, G.J., and Plasterk, R.H. (2003). A micrococcal nuclease homologue in RNAi effector complexes. *Nature* 425, 411-414.
- Caudy, A.A., Myers, M., Hannon, G.J., and Hammond, S.M. (2002). Fragile X-related protein and VIG associate with the RNA interference machinery. *Genes Dev* 16, 2491-2496.
- Chang, K.-Y., and Ramos, A. (2005). The double-stranded RNA-binding motif, a versatile macromolecular docking platform. *FEBS J* 272, 2109-2117.
- Chen, P.Y., Weinmann, L., Gaidatzis, D., Pei, Y., Zavolan, M., Tuschl, T., and Meister, G. (2008). Strand-specific 5'-O-methylation of siRNA duplexes controls guide strand selection and targeting specificity. *RNA* 14, 263-274.
- Chendrimada, T.P., Gregory, R.I., Kumaraswamy, E., Norman, J., Cooch, N., Nishikura, K., and Shiekhattar, R. (2005). TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature* 436, 740-744.
- Chiu, Y., and Rana, T.M. (2003). siRNA function in RNAi: A chemical modification analysis. *RNA* 9, 1034-1048.

- Clamme, J.P., Azoulay, J., and Mely, Y. (2003). Monitoring of the formation and dissociation of polyethylenimine/DNA complexes by two photon fluorescence correlation spectroscopy. *Biophys J* 84, 1960-1968.
- Conklin, D.S. (2003). RNA-Interference-Based Silencing of Mammalian Gene Expression. *ChemBioChem* 4, 1033-1039.
- Consortium (2004). Finishing the euchromatic sequence of the human genome. *Nature* 431, 931-945.
- Corey, D.R. (2007). Chemical modification: the key to clinical application of RNA interference? *J Clin Invest* 117, 3615-3622.
- Cosentino, G.P., Venkatesan, S., Serluca, F.C., Green, S.R., Mathews, M.B., and Sonenberg, N. (1995). Double-stranded-RNA-dependent protein kinase and TAR RNA-binding protein form homo- and heterodimers in vivo. *Proc Natl Acad Sci U S A* 92, 9445-9449.
- Cox, K.M., Sterling, J.D., Regan, J.T., Gasdaska, J.R., Frantz, K.K., Peele, C.G., Black, A., Passmore, D., Moldovan-Loomis, C., Srinivasan, M., *et al.* (2006). Glycan optimization of a human monoclonal antibody in the aquatic plant *Lemna minor*. *Nat Biotechnol* 24, 1591-1597.
- Daher, A., Longuet, M., Dorin, D., Bois, F., Segéral, E., Bannwarth, S., Battisti, P.L., Purcell, D.F., Benarous, R., Vaquero, C., *et al.* (2001). Two dimerization domains in the trans-activation response RNA-binding protein (TRBP) individually reverse the protein kinase R inhibition of HIV-1 long terminal repeat expression. *J Biol Chem* 276, 33899-33905.
- Daviet, L., Erard, M., Dorin, D., Duarte, M., Vaquero, C., and Gatignol, A. (2000). Analysis of a binding difference between the two dsRNA-binding domains in TRBP reveals the modular function of a KR-helix motif. *Eur J Biochem* 267, 2419-2431.
- de Fougerolles, A., Vornlocher, H.P., Maraganore, J., and Lieberman, J. (2007). Interfering with disease: a progress report on siRNA-based therapeutics. *Nat Rev Drug Discov* 6, 443-453.
- de Fougerolles, A.R. (2008). Delivery vehicles for small interfering RNA in vivo. *Hum Gene Ther* 19, 125-132.
- de Martimprey, H., Vauthier, C., Malvy, C., and Couvreur, P. (2009). Polymer nanocarriers for the delivery of small fragments of nucleic acids: oligonucleotides and siRNA. *Eur J Pharm Biopharm* 71, 490-504.
- Ding, Y., Chan, C.Y., and Lawrence, C.E. (2004). Sfold web server for statistical folding and rational design of nucleic acids. *Nucleic Acids Res* 32, W135-141.

- Dodo, H.W., Konan, K.N., Chen, F.C., Egnin, M., and Viquez, O.M. (2008). Alleviating peanut allergy using genetic engineering: the silencing of the immunodominant allergen Ara h 2 leads to its significant reduction and a decrease in peanut allergenicity. *Plant Biotechnol J* 6, 135-145.
- Doench, J.G., and Sharp, P.A. (2004). Specificity of microRNA target selection in translational repression. *Genes Dev* 18, 504-511.
- Donker, R.B., Mouillet, J.-F., Nelson, D.M., and Sadovsky, Y. (2007). The expression of Argonaute2 and related microRNA biogenesis proteins in normal and hypoxic trophoblasts. *Mol Hum Reprod* 13, 273-279.
- Doshi, K., Cannone, J., Cobaugh, C., and Gutell, R. (2004). Evaluation of the suitability of free-energy minimization using nearest-neighbor energy parameters for RNA secondary structure prediction. *BMC Bioinform* 5, 105.
- Douglas, K.L., Piccirillo, C.A., and Tabrizian, M. (2008). Cell line-dependent internalization pathways and intracellular trafficking determine transfection efficiency of nanoparticle vectors. *Eur J Pharm Biopharm* 68, 676-687.
- Doyle, M., and Jantsch, M.F. (2002). New and old roles of the double-stranded RNA-binding domain. *J Struct Biol* 140, 147-153.
- Eckmann, C.R., and Jantsch, M.F. (1997). Xlrpba, a Double-stranded RNA-binding Protein Associated with Ribosomes and Heterogeneous Nuclear RNPs. *J Cell Biol* 138, 239-253.
- Elbashir, S.M., Harborth, J., Weber, K., and Tuschl, T. (2002). Analysis of gene function in somatic mammalian cells using small interfering RNAs. *Methods* 26, 199-213.
- Elbashir, S.M., Lendeckel, W., and Tuschl, T. (2001a). RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev* 15, 188-200.
- Elbashir, S.M., Martinez, J., Patkaniowska, A., Lendeckel, W., and Tuschl, T. (2001b). Functional anatomy of siRNAs for mediating efficient RNAi in *Drosophila melanogaster* embryo lysate. *EMBO J* 20, 6877-6888.
- Elmen, J., Thonberg, H., Ljungberg, K., Frieden, M., Westergaard, M., Xu, Y., Wahren, B., Liang, Z., Orum, H., Koch, T., *et al.* (2005). Locked nucleic acid (LNA) mediated improvements in siRNA stability and functionality. *Nucleic Acids Res* 33, 439-447.
- Erard, M., Barker, D.G., Amalric, F., Jeang, K.-T., and Gatignol, A. (1998). An Arg/Lys-rich core peptide mimics TRBP binding to the HIV-1 TAR RNA upper-stem/loop. *J Mol Biol* 279, 1085-1099.

- Far, R., and Sczakiel, G. (2003). The activity of siRNA in mammalian cells is related to structural target accessibility: a comparison with antisense oligonucleotides. *Nucleic Acids Res* 31, 4417-4424.
- Fedorov, Y., Anderson, E.M., Birmingham, A., Reynolds, A., Karpilow, J., Robinson, K., Leake, D., Marshall, W.S., and Khvorova, A. (2006). Off-target effects by siRNA can induce toxic phenotype. *RNA* 12, 1188-1196.
- Felgner, P.L., Gadek, T.R., Holm, M., Roman, R., Chan, H.W., Wenz, M., Northrop, J.P., Ringold, G.M., and Danielsen, M. (1987). Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure. *Proc Natl Acad Sci U S A* 84, 7413-7417.
- Fierro-Monti, I., and Mathews, M.B. (2000). Proteins binding to duplexed RNA: one motif, multiple functions. *Trends Biochem Sci* 25, 241-246.
- Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., and Mello, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391, 806-811.
- Forstemann, K., Tomari, Y., Du, T., Vagin, V.V., Denli, A.M., Bratu, D.P., Klattenhoff, C., Theurkauf, W.E., and Zamore, P.D. (2005). Normal microRNA maturation and germline stem cell maintenance requires Loquacious, a double-stranded RNA-binding domain protein. *PLoS Biol* 3, e236.
- Garcia, M.A., Gil, J., Ventoso, I., Guerra, S., Domingo, E., Rivas, C., and Esteban, M. (2006). Impact of Protein Kinase PKR in Cell Biology: from Antiviral to Antiproliferative Action. *Microbiol Mol Biol Rev* 70, 1032-1060.
- Gatignol, A., Buckler-White, A., Berkhout, B., and Jeang, K.T. (1991). Characterization of a human TAR RNA-binding protein that activates the HIV-1 LTR. *Science* 251, 1597-1600.
- Gatignol, A., Buckler, C., and Jeang, K.T. (1993). Relatedness of an RNA-binding motif in human immunodeficiency virus type 1 TAR RNA-binding protein TRBP to human P1/dsI kinase and *Drosophila* staufen. *Mol Cell Biol* 13, 2193-2202.
- Ge, G., Wong, G.W., and Luo, B. (2005). Prediction of siRNA knockdown efficiency using artificial neural network models. *Biochem Biophys Res Commun* 336, 723-728.
- Geng, C.M., and Ding, H.L. (2008). Double-mismatched siRNAs enhance selective gene silencing of a mutant ALS-causing allele. *Acta pharmacologica Sinica* 29, 211-216.
- Goutelle, S., Maurin, M., Rougier, F., Barbaut, X., Bourguignon, L., Ducher, M., and Maire, P. (2008). The Hill equation: a review of its capabilities in pharmacological modelling. *Fundamental & clinical pharmacology* 22, 633-648.

- Gredell, J.A., Berger, A.K., and Walton, S.P. (2008). Impact of target mRNA structure on siRNA silencing efficiency: A large-scale study. *Biotechnol Bioeng* 100, 744-755.
- Gregory, R.I., Chendrimada, T.P., Cooch, N., and Shiekhattar, R. (2005). Human RISC Couples MicroRNA Biogenesis and Posttranscriptional Gene Silencing. *Cell* 123, 631-640.
- Grimm, D., Streetz, K.L., Jopling, C.L., Storm, T.A., Pandey, K., Davis, C.R., Marion, P., Salazar, F., and Kay, M.A. (2006). Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways. *Nature* 441, 537-541.
- Haase, A.D., Jaskiewicz, L., Zhang, H., Laine, S., Sack, R., Gatignol, A., and Filipowicz, W. (2005). TRBP, a regulator of cellular PKR and HIV-1 virus expression, interacts with Dicer and functions in RNA silencing. *EMBO Rep* 6, 961-967.
- Haley, B., and Zamore, P.D. (2004). Kinetic analysis of the RNAi enzyme complex. *Nat Struct Mol Biol* 11, 599-606.
- Hannon, G.J. (2002). RNA interference. *Nature* 418, 244-251.
- Hargittai, M.R., Gorelick, R.J., Rouzina, I., and Musier-Forsyth, K. (2004). Mechanistic insights into the kinetics of HIV-1 nucleocapsid protein-facilitated tRNA annealing to the primer binding site. *J Mol Biol* 337, 951-968.
- Heale, B.S., Soifer, H.S., Bowers, C., and Rossi, J.J. (2005). siRNA target site secondary structure predictions using local stable substructures. *Nucleic Acids Res* 33, e30.
- Hein, C.D., Liu, X.M., and Wang, D. (2008). Click chemistry, a powerful tool for pharmaceutical sciences. *Pharm Res* 25, 2216-2230.
- Hofacker, I.L. (2003). Vienna RNA secondary structure server. *Nucleic Acids Res* 31, 3429-3431.
- Hofacker, I.L., Fontana, W., Stadler, P.F., Bonhoeffer, L.S., Tacker, M., and Schuster, P. (1994). Fast Folding and Comparison of Rna Secondary Structures. *Monatshefte Fur Chemie* 125, 167-188.
- Hohjoh, H. (2002). RNA interference (RNAi) induction with various types of synthetic oligonucleotide duplexes in cultured human cells. *FEBS Lett* 521, 195-199.
- Hornung, V., Guenther-Biller, M., Bourquin, C., Ablasser, A., Schlee, M., Uematsu, S., Noronha, A., Manoharan, M., Akira, S., de Fougerolles, A., *et al.* (2005). Sequence-specific potent induction of IFN-alpha by short interfering RNA in plasmacytoid dendritic cells through TLR7. *Nat Med* 11, 263-270.
- Hu-Lieskovan, S., Heidel, J.D., Bartlett, D.W., Davis, M.E., and Triche, T.J. (2005). Sequence-specific knockdown of EWS-FLI1 by targeted, nonviral delivery of small

interfering RNA inhibits tumor growth in a murine model of metastatic Ewing's sarcoma. *Cancer Res* 65, 8984-8992.

Hutvagner, G. (2005). Small RNA asymmetry in RNAi: Function in RISC assembly and gene regulation. *FEBS Letters* 579, 5850-5857.

Hutvagner, G., and Simard, M.J. (2008). Argonaute proteins: key players in RNA silencing. *Nat Rev Mol Cell Biol* 9, 22-32.

Hutvagner, G., and Zamore, P.D. (2002). A microRNA in a multiple-turnover RNAi enzyme complex. *Science* 297, 2056-2060.

Ishizuka, A., Siomi, M.C., and Siomi, H. (2002). A Drosophila fragile X protein interacts with components of RNAi and ribosomal proteins. *Genes Dev* 16, 2497-2508.

Jackson, A.L., Bartz, S.R., Schelter, J., Kobayashi, S.V., Burchard, J., Mao, M., Li, B., Cavet, G., and Linsley, P.S. (2003). Expression profiling reveals off-target gene regulation by RNAi. *Nat Biotechnol* 21, 635-637.

Jackson, A.L., Burchard, J., Leake, D., Reynolds, A., Schelter, J., Guo, J., Johnson, J.M., Lim, L., Karpilow, J., Nichols, K., *et al.* (2006). Position-specific chemical modification of siRNAs reduces "off-target" transcript silencing. *RNA* 12, 1197-1205.

Jagla, B., Aulner, N., Kelly, P.D., Song, D.A., Volchuk, A., Zatorski, A., Shum, D., Mayer, T., De Angelis, D.A., Ouerfelli, O., *et al.* (2005). Sequence characteristics of functional siRNAs. *RNA* 11, 864-872.

Jakymiw, A., Lian, S., Eystathioy, T., Li, S., Satoh, M., Hamel, J.C., Fritzler, M.J., and Chan, E.K. (2005). Disruption of GW bodies impairs mammalian RNA interference. *Nat Cell Biol* 7, 1267-1274.

Jeong, J.H., and Park, T.G. (2002). Poly(L-lysine)-g-poly(D,L-lactic-co-glycolic acid) micelles for low cytotoxic biodegradable gene delivery carriers. *J Control Release* 82, 159-166.

Jiang, X., Vogel, E.B., Smith, M.R., and Baker, G.L. (2008). "Clickable" Polyglycolides: Tunable Synthons for Thermoresponsive, Degradable Polymers. *Macromolecules* 41, 1937-1944.

Judge, A., and Maclachlan, I. (2008). Overcoming the innate immune response to small interfering RNA. *Hum Gene Ther* 19, 111-124.

Judge, A.D., Sood, V., Shaw, J.R., Fang, D., McClintock, K., and Maclachlan, I. (2005). Sequence-dependent stimulation of the mammalian innate immune response by synthetic siRNA. *Nat Biotechnol* 23, 457-462.

- Kariko, K., Bhuyan, P., Capodici, J., and Weissman, D. (2004). Small Interfering RNAs Mediate Sequence-Independent Gene Suppression and Induce Immune Activation by Signaling through Toll-Like Receptor 3. *J Immunol* 172, 6545-6549.
- Katoh, T., and Suzuki, T. (2007). Specific residues at every third position of siRNA shape its efficient RNAi activity. *Nucleic Acids Res*, e27.
- Kawahara, Y., Zinshteyn, B., Chendrimada, T.P., Shiekhattar, R., and Nishikura, K. (2007). RNA editing of the microRNA-151 precursor blocks cleavage by the Dicer-TRBP complex. *EMBO Rep*.
- Khvorova, A., Reynolds, A., and Jayasena, S.D. (2003). Functional siRNAs and miRNAs exhibit strand bias. *Cell* 115, 209-216.
- Kim, D.H., Behlke, M.A., Rose, S.D., Chang, M.S., Choi, S., and Rossi, J.J. (2005). Synthetic dsRNA Dicer substrates enhance RNAi potency and efficacy. *Nat Biotechnol* 23, 222-226.
- Kim, D.H., Longo, M., Han, Y., Lundberg, P., Cantin, E., and Rossi, J.J. (2004). Interferon induction by siRNAs and ssRNAs synthesized by phage polymerase. *Nat Biotechnol* 22, 321-325.
- Kim, D.H., and Rossi, J.J. (2003). Coupling of RNAi-mediated target downregulation with gene replacement. *Antisense Nucleic Acid Drug Dev* 13, 151-155.
- Kim, E.J., Kramer, S.F., Hebert, C.G., Valdes, J.J., and Bentley, W.E. (2007). Metabolic engineering of the baculovirus-expression system via inverse "shotgun" genomic analysis and RNA interference (dsRNA) increases product yield and cell longevity. *Biotechnol Bioeng* 98, 645-654.
- Kim, I., Liu, C.W., and Puglisi, J.D. (2006). Specific Recognition of HIV TAR RNA by the dsRNA Binding Domains (dsRBD1-dsRBD2) of PKR. *J Mol Biol* 358, 430-442.
- Kini, H.K., and Walton, S.P. (2007). In vitro binding of single-stranded RNA by human Dicer. *FEBS Lett* 581, 5611-5616.
- Kircheis, R., Wightman, L., and Wagner, E. (2001). Design and gene delivery activity of modified polyethylenimines. *Adv Drug Deliv Rev* 53, 341-358.
- Kleinman, M.E., Yamada, K., Takeda, A., Chandrasekaran, V., Nozaki, M., Baffi, J.Z., Albuquerque, R.J., Yamasaki, S., Itaya, M., Pan, Y., *et al.* (2008). Sequence- and target-independent angiogenesis suppression by siRNA via TLR3. *Nature* 452, 591-597.
- Koberle, C., Kaufmann, S.H., and Patzel, V. (2006). Selecting effective siRNAs based on guide RNA structure. *Nat Protoc* 1, 1832-1839.

- Kok, K.H., Ng, M.H., Ching, Y.P., and Jin, D.Y. (2007). Human TRBP and PACT Directly Interact with Each Other and Associate with Dicer to Facilitate the Production of Small Interfering RNA. *J Biol Chem* *282*, 17649-17657.
- Kolb, H., Finn, M., and Sharpless, K. (2001). Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angew Chem Int Ed Engl* *40*, 2004-2021.
- Kunath, K., von Harpe, A., Petersen, H., Fischer, D., Voigt, K., Kissel, T., and Bickel, U. (2002). The structure of PEG-modified poly(ethylene imines) influences biodistribution and pharmacokinetics of their complexes with NF-kappaB decoy in mice. *Pharm Res* *19*, 810-817.
- Lai, E.C. (2002). Micro RNAs are complementary to 3' UTR sequence motifs that mediate negative post-transcriptional regulation. *Nat Genet* *30*, 363-364.
- Lamberton, J.S., and Christian, A.T. (2003). Varying the nucleic acid composition of siRNA molecules dramatically varies the duration and degree of gene silencing. *Mol Biotechnol* *24*, 111-120.
- Laraki, G., Clerzius, G., Daher, A., Melendez-Pena, C., Daniels, S., and Gagnon, A. (2008). Interactions between the double-stranded RNA-binding proteins TRBP and PACT define the Medial domain that mediates protein-protein interactions. *RNA Biol* *5*, 92-103.
- Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Radmark, O., Kim, S., *et al.* (2003). The nuclear RNase III Drosha initiates microRNA processing. *Nature* *425*, 415-419.
- Lee, Y., Hur, I., Park, S.Y., Kim, Y.K., Suh, M.R., and Kim, V.N. (2006). The role of PACT in the RNA silencing pathway. *EMBO J* *25*, 522-532.
- Leonard, J.N., Shah, P.S., Burnett, J.C., and Schaffer, D.V. (2008). HIV Evades RNA Interference Directed at TAR by an Indirect Compensatory Mechanism. *Cell Host Microbe* *4*, 484-494.
- Li, S., Peters, G.A., Ding, K., Zhang, X., Qin, J., and Sen, G.C. (2006). Molecular basis for PKR activation by PACT or dsRNA. *Proc Natl Acad Sci U S A* *103*, 10005-10010.
- Lian, S., Fritzler, M.J., Katz, J., Hamazaki, T., Terada, N., Satoh, M., and Chan, E.K. (2007). Small interfering RNA-mediated silencing induces target-dependent assembly of GW/P bodies. *Mol Biol Cell* *18*, 3375-3387.
- Lima, W.F., Murray, H., Nichols, J.G., Wu, H., Sun, H., Prakash, T.P., Berdeja, A.R., Gaus, H.J., and Croke, S.T. (2009). Human Dicer Binds Short Single-strand and Double-strand RNA with High Affinity and Interacts with Different Regions of the Nucleic Acids. *J Biol Chem* *284*, 2535-2548.

- Liu, J., Carmell, M.A., Rivas, F.V., Marsden, C.G., Thomson, J.M., Song, J.J., Hammond, S.M., Joshua-Tor, L., and Hannon, G.J. (2004). Argonaute2 Is the Catalytic Engine of Mammalian RNAi. *Science* 305, 1437-1441.
- Liu, J., Rivas, F.V., Wohlschlegel, J., Yates, J.R., Parker, R., and Hannon, G.J. (2005a). A role for the P-body component GW182 in microRNA function. *Nat Cell Biol* 7, 1261-1266.
- Liu, J., Valencia-Sanchez, M.A., Hannon, G.J., and Parker, R. (2005b). MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies. *Nat Cell Biol* 7, 719-723.
- Liu, Q., Rand, T.A., Kalidas, S., Du, F., Kim, H.E., Smith, D.P., and Wang, X. (2003). R2D2, a bridge between the initiation and effector steps of the Drosophila RNAi pathway. *Science* 301, 1921-1925.
- Long, D., Lee, R., Williams, P., Chan, C.Y., Ambros, V., and Ding, Y. (2007). Potent effect of target structure on microRNA function. *Nat Struct Mol Biol* 14, 287.
- Lu, Z.J., and Mathews, D.H. (2007). Efficient siRNA selection using hybridization thermodynamics. *Nucleic Acids Res* 36, 640-647.
- Lund, E., Guttinger, S., Calado, A., Dahlberg, J.E., and Kutay, U. (2004). Nuclear Export of MicroRNA Precursors. *Science* 303, 95-98.
- Luo, K.Q., and Chang, D.C. (2004). The gene-silencing efficiency of siRNA is strongly dependent on the local structure of mRNA at the targeted region. *Biochem Biophys Res Commun* 318, 303-310.
- Ma, J.B., Ye, K., and Patel, D.J. (2004). Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. *Nature* 429, 318-322.
- Ma, J.B., Yuan, Y.R., Meister, G., Pei, Y., Tuschl, T., and Patel, D.J. (2005). Structural basis for 5'-end-specific recognition of guide RNA by the *A. fulgidus* Piwi protein. *Nature* 434, 666-670.
- MacRae, I.J., Ma, E., Zhou, M., Robinson, C.V., and Doudna, J.A. (2008). In vitro reconstitution of the human RISC-loading complex. *Proc Natl Acad Sci U S A* 105, 512-517.
- MacRae, I.J., Zhou, K., and Doudna, J.A. (2007). Structural determinants of RNA recognition and cleavage by Dicer. *Nat Struct Mol Biol* 14, 934-940.
- MacRae, I.J., Zhou, K., Li, F., Repic, A., Brooks, A.N., Cande, W.Z., Adams, P.D., and Doudna, J.A. (2006). Structural basis for double-stranded RNA processing by Dicer. *Science* 311, 195-198.

- Manche, L., Green, S.R., Schmedt, C., and Mathews, M.B. (1992). Interactions between double-stranded RNA regulators and the protein kinase DAI. *Mol Cell Biol* *12*, 5238-5248.
- Maniataki, E., and Mourelatos, Z. (2005). A human, ATP-independent, RISC assembly machine fueled by pre-miRNA. *Genes Dev* *19*, 2979-2990.
- Manjunath, N., Wu, H., Subramanya, S., and Shankar, P. (2009). Lentiviral delivery of short hairpin RNAs. *Advanced Drug Delivery Reviews* *61*, 732-745.
- Mann, A., Richa, R., and Ganguli, M. (2008). DNA condensation by poly-L-lysine at the single molecule level: Role of DNA concentration and polymer length. *J Control Release* *125*, 252-262.
- Mao, S., Neu, M., Gernershaus, O., Merkel, O., Sitterberg, J., Bakowsky, U., and Kissel, T. (2006). Influence of Polyethylene Glycol Chain Length on the Physicochemical and Biological Properties of Poly(ethylene imine)-*graft*-Poly(ethylene glycol) Block Copolymer/SiRNA Polyplexes. *Bioconjug Chem* *17*, 1209-1218.
- Markham, N.R., and Zuker, M. (2005). DINAMelt web server for nucleic acid melting prediction. *Nucleic Acids Res* *33*, W577-581.
- Markham, N.R., and Zuker, M. (2008). UNAFold: software for nucleic acid folding and hybridization. *Methods Mol Biol* *453*, 3-31.
- Martinez, J., and Tuschl, T. (2004). RISC is a 5' phosphomonoester-producing RNA endonuclease. *Genes Dev* *18*, 975-980.
- Mathews, D.H., Disney, M.D., Childs, J.L., Schroeder, S.J., Zuker, M., and Turner, D.H. (2004). Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. *Proc Natl Acad Sci U S A* *101*, 7287-7292.
- Mathews, D.H., Sabina, J., Zuker, M., and Turner, D.H. (1999). Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J Mol Biol* *288*, 911.
- Matranga, C., Tomari, Y., Shin, C., Bartel, D.P., and Zamore, P.D. (2005). Passenger-Strand Cleavage Facilitates Assembly of siRNA into Ago2-Containing RNAi Enzyme Complexes. *Cell* *123*, 607-620.
- Matveeva, O., Nechipurenko, Y., Rossi, L., Moore, B., Saetrom, P., Ogurtsov, A.Y., Atkins, J.F., and Shabalina, S.A. (2007). Comparison of approaches for rational siRNA design leading to a new efficient and transparent method. *Nucleic Acids Res* *35*, e63.
- McBride, J.L., Boudreau, R.L., Harper, S.Q., Staber, P.D., Monteys, A.M., Martins, I., Gilmore, B.L., Burstein, H., Peluso, R.W., Polisky, B., *et al.* (2008). Artificial miRNAs

mitigate shRNA-mediated toxicity in the brain: implications for the therapeutic development of RNAi. *Proc Natl Acad Sci U S A* 105, 5868-5873.

Meister, G., Landthaler, M., Patkaniowska, A., Dorsett, Y., Teng, G., and Tuschl, T. (2004). Human Argonaute2 Mediates RNA Cleavage Targeted by miRNAs and siRNAs. *Mol Cell* 15, 185-197.

Meister, G., Landthaler, M., Peters, L., Chen, P.Y., Urlaub, H., Luhrmann, R., and Tuschl, T. (2005). Identification of Novel Argonaute-Associated Proteins. *Curr Biol* 15, 2149-2155.

Mello, C.C., and Conte, D., Jr. (2004). Revealing the world of RNA interference. *Nature* 431, 338-342.

Meylan, E., and Tschopp, J. (2006). Toll-like receptors and RNA helicases: two parallel ways to trigger antiviral responses. *Mol Cell* 22, 561-569.

Muckstein, U., Tafer, H., Hackermuller, J., Bernhart, S.H., Stadler, P.F., and Hofacker, I.L. (2006). Thermodynamics of RNA-RNA binding. *Bioinformatics* 22, 1177-1182.

Murphy, D., Dancis, B., and Brown, J.R. (2008). The evolution of core proteins involved in microRNA biogenesis. *BMC Evol Biol* 8, 92.

Nallagatla, S.R., and Bevilacqua, P.C. (2008). Nucleoside modifications modulate activation of the protein kinase PKR in an RNA structure-specific manner. *RNA* 14, 1201-1213.

Nallagatla, S.R., Hwang, J., Toroney, R., Zheng, X., Cameron, C.E., and Bevilacqua, P.C. (2007). 5'-triphosphate-dependent activation of PKR by RNAs with short stem-loops. *Science* 318, 1455-1458.

Nanduri, S., Carpick, B.W., Yang, Y., Williams, B.R., and Qin, J. (1998). Structure of the double-stranded RNA-binding domain of the protein kinase PKR reveals the molecular basis of its dsRNA-mediated activation. *EMBO J* 17, 5458-5465.

Nielsen, C.B., Shomron, N., Sandberg, R., Hornstein, E., Kitzman, J., and Burge, C.B. (2007). Determinants of targeting by endogenous and exogenous microRNAs and siRNAs. *RNA* 13, 1894-1910.

Okamura, K., Ishizuka, A., Siomi, H., and Siomi, M.C. (2004). Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways. *Genes Dev* 18, 1655-1666.

Overhoff, M., Alken, M., Far, R.K., Lemaitre, M., Lebleu, B., Sczakiel, G., and Robbins, I. (2005). Local RNA target structure influences siRNA efficacy: a systematic global analysis. *J Mol Biol* 348, 871-881.

- Pardridge, W.M. (2007). shRNA and siRNA delivery to the brain. *Adv Drug Deliv Rev* 59, 141-152.
- Parker, G.S., Maity, T.S., and Bass, B.L. (2008). dsRNA binding properties of RDE-4 and TRBP reflect their distinct roles in RNAi. *J Mol Biol* 384, 967-979.
- Patel, R.C., Stanton, P., McMillan, N.M., Williams, B.R., and Sen, G.C. (1995). The interferon-inducible double-stranded RNA-activated protein kinase self-associates in vitro and in vivo. *Proc Natl Acad Sci U S A* 92, 8283-8287.
- Patzel, V., Rutz, S., Dietrich, I., Koberle, C., Scheffold, A., and Kaufmann, S.H. (2005). Design of siRNAs producing unstructured guide-RNAs results in improved RNA interference efficiency. *Nat Biotechnol* 23, 1440-1444.
- Pellino, J.L., Jaskiewicz, L., Filipowicz, W., and Sontheimer, E.J. (2005). ATP modulates siRNA interactions with an endogenous human Dicer complex. *RNA* 11, 1719-1724.
- Provost, P., Dishart, D., Doucet, J., Frendewey, D., Samuelsson, B., and Radmark, O. (2002). Ribonuclease activity and RNA binding of recombinant human Dicer. *EMBO J* 21, 5864-5874.
- Pun, S.H., Tack, F., Bellocq, N.C., Cheng, J., Grubbs, B.H., Jensen, G.S., Davis, M.E., Brewster, M., Janicot, M., Janssens, B., *et al.* (2004). Targeted delivery of RNA-cleaving DNA enzyme (DNAzyme) to tumor tissue by transferrin-modified, cyclodextrin-based particles. *Cancer Biol Ther* 3, 641-650.
- Ramos, A., Grunert, S., Adams, J., Micklem, D.R., Proctor, M.R., Freund, S., Bycroft, M., St Johnston, D., and Varani, G. (2000). RNA recognition by a Staufien double-stranded RNA-binding domain. *EMBO J* 19, 997-1009.
- Rana, T.M. (2007). Illuminating the silence: understanding the structure and function of small RNAs. *Nat Rev Mol Cell Biol* 8, 23-36.
- Rana, T.M., and Jeang, K.T. (1999). Biochemical and functional interactions between HIV-1 Tat protein and TAR RNA. *Arch Biochem Biophys* 365, 175-185.
- Rand, T.A., Petersen, S., Du, F., and Wang, X. (2005). Argonaute2 Cleaves the Anti-Guide Strand of siRNA during RISC Activation. *Cell* 123, 621-629.
- Reynolds, A., Anderson, E.M., Vermeulen, A., Fedorov, Y., Robinson, K., Leake, D., Karpilow, J., Marshall, W.S., and Khvorova, A. (2006). Induction of the interferon response by siRNA is cell type- and duplex length-dependent. *RNA* 12, 988-993.
- Reynolds, A., Leake, D., Boese, Q., Scaringe, S., Marshall, W.S., and Khvorova, A. (2004). Rational siRNA design for RNA interference. *Nat Biotechnol* 22, 326-330.

- Rivas, F.V., Tolia, N.H., Song, J.J., Aragon, J.P., Liu, J., Hannon, G.J., and Joshua-Tor, L. (2005). Purified Argonaute2 and an siRNA form recombinant human RISC. *Nat Struct Mol Biol* *12*, 340-349.
- Robb, G.B., and Rana, T.M. (2007). RNA Helicase A Interacts with RISC in Human Cells and Functions in RISC Loading. *Mol Cell* *26*, 523-537.
- Rose, S.D., Kim, D.H., Amarzguioui, M., Heidel, J.D., Collingwood, M.A., Davis, M.E., Rossi, J.J., and Behlke, M.A. (2005). Functional polarity is introduced by Dicer processing of short substrate RNAs. *Nucleic Acids Res* *33*, 4140-4156.
- Rossi, J.J. (2008). Expression strategies for short hairpin RNA interference triggers. *Hum Gene Ther* *19*, 313-317.
- Rozema, D.B., Ekena, K., Lewis, D.L., Loomis, A.G., and Wolff, J.A. (2003). Endosomolysis by Masking of a Membrane-Active Agent (EMMA) for Cytoplasmic Release of Macromolecules. *Bioconjug Chem* *14*, 51-57.
- Rozema, D.B., Lewis, D.L., Wakefield, D.H., Wong, S.C., Klein, J.J., Roesch, P.L., Bertin, S.L., Reppen, T.W., Chu, Q., Blokhin, A.V., *et al.* (2007). Dynamic PolyConjugates for targeted in vivo delivery of siRNA to hepatocytes. *Proc Natl Acad Sci U S A* *104*, 12982-12987.
- Ryter, J.M., and Schultz, S.C. (1998). Molecular basis of double-stranded RNA-protein interactions: structure of a dsRNA-binding domain complexed with dsRNA. *EMBO J* *17*, 7505-7513.
- Saunders, L.R., and Barber, G.N. (2003). The dsRNA binding protein family: critical roles, diverse cellular functions. *Faseb J* *17*, 961-983.
- Scacheri, P.C., Rozenblatt-Rosen, O., Caplen, N.J., Wolfsberg, T.G., Umayam, L., Lee, J.C., Hughes, C.M., Shanmugam, K.S., Bhattacharjee, A., Meyerson, M., *et al.* (2004). Short interfering RNAs can induce unexpected and divergent changes in the levels of untargeted proteins in mammalian cells. *PNAS* *101*, 1892-1897.
- Schubert, S., Grunweller, A., Erdmann, V.A., and Kurreck, J. (2005). Local RNA Target Structure Influences siRNA Efficacy: Systematic Analysis of Intentionally Designed Binding Regions. *J Mol Biol* *348*, 883-893.
- Schwarz, D.S., Ding, H., Kennington, L., Moore, J.T., Schelter, J., Burchard, J., Linsley, P.S., Aronin, N., Xu, Z., and Zamore, P.D. (2006). Designing siRNA that distinguish between genes that differ by a single nucleotide. *PLoS Genet* *2*, e140.
- Schwarz, D.S., Hutvagner, G., Du, T., Xu, Z., Aronin, N., and Zamore, P.D. (2003). Asymmetry in the assembly of the RNAi enzyme complex. *Cell* *115*, 199-208.

- Seeman, N.C., Rosenberg, J.M., and Rich, A. (1976). Sequence-specific recognition of double helical nucleic acids by proteins. *Proc Natl Acad Sci U S A* 73, 804-808.
- Shabalina, S.A., Spiridonov, A.N., and Ogurtsov, A.Y. (2006). Computational models with thermodynamic and composition features improve siRNA design. *BMC Bioinformatics* 7, 65.
- Shah, J.K., Garner, H.R., White, M.A., Shames, D.S., and Minna, J.D. (2007). siR: siRNA Information Resource, a web-based tool for siRNA sequence design and analysis and an open access siRNA database. *BMC Bioinform* 8, 178.
- Shao, Y., Chan, C.Y., Maliyekkel, A., Lawrence, C.E., Roninson, I.B., and Ding, Y. (2007). Effect of target secondary structure on RNAi efficiency. *RNA* 13, 1631-1640.
- Sharpless, K.B., and Kolb, H.C. (1999). Click chemistry. A concept for merging process and discovery chemistry. *Abstracts of Papers of the American Chemical Society* 217, U95-U95.
- Siolas, D., Lerner, C., Burchard, J., Ge, W., Linsley, P.S., Paddison, P.J., Hannon, G.J., and Cleary, M.A. (2005). Synthetic shRNAs as potent RNAi triggers. *Nat Biotechnol* 23, 227-231.
- Siomi, H., and Siomi, M.C. (2009). On the road to reading the RNA-interference code. *Nature* 457, 396-404.
- Sioud, M. (2006). Innate sensing of self and non-self RNAs by Toll-like receptors. *Trends Mol Med* 12, 167-176.
- Sipa, K., Sochacka, E., Kazmierczak-Baranska, J., Maszewska, M., Janicka, M., Nowak, G., and Nawrot, B. (2007). Effect of base modifications on structure, thermodynamic stability, and gene silencing activity of short interfering RNA. *RNA* 13, 1301-1316.
- Sledz, C.A., Holko, M., de Veer, M.J., Silverman, R.H., and Williams, B.R. (2003). Activation of the interferon system by short-interfering RNAs. *Nat Cell Biol* 5, 834-839.
- Soifer, H.S., Sano, M., Sakurai, K., Chomchan, P., Saetrom, P., Sherman, M.A., Collingwood, M.A., Behlke, M.A., and Rossi, J.J. (2008). A role for the Dicer helicase domain in the processing of thermodynamically unstable hairpin RNAs. *Nucleic Acids Res* 36, 6511-6522.
- Song, J.J., Smith, S.K., Hannon, G.J., and Joshua-Tor, L. (2004). Crystal Structure of Argonaute and Its Implications for RISC Slicer Activity. *Science* 305, 1434-1437.
- Sontheimer, E.J. (1994). Site-specific RNA crosslinking with 4-thiouridine. *Mol Biol Rep* 20, 35-44.

- Stein, C.A. (1998). How to design an antisense oligodeoxynucleotide experiment: a consensus approach. *Antisense Nucleic Acid Drug Dev* 8, 129-132.
- Swami, A., Kurupati, R.K., Pathak, A., Singh, Y., Kumar, P., and Gupta, K.C. (2007). A unique and highly efficient non-viral DNA/siRNA delivery system based on PEI-bisepoxide nanoparticles. *Biochem Biophys Res Commun* 362, 835-841.
- Tahbaz, N., Kolb, F.A., Zhang, H., Jaronczyk, K., Filipowicz, W., and Hobman, T.C. (2004). Characterization of the interactions between mammalian PAZ PIWI domain proteins and Dicer. *EMBO Rep* 5, 189-194.
- Takashima, Y., Saito, R., Nakajima, A., Oda, M., Kimura, A., Kanazawa, T., and Okada, H. (2007). Spray-drying preparation of microparticles containing cationic PLGA nanospheres as gene carriers for avoiding aggregation of nanospheres. *Int J Pharm* 343, 262-269.
- Tang, G.L., Galili, G., and Zhuang, X. (2007). RNAi and microRNA: breakthrough technologies for the improvement of plant nutritional value and metabolic engineering. *Metabolomics* 3, 357-369.
- Terrazas, M., and Kool, E.T. (2009). RNA major groove modifications improve siRNA stability and biological activity. *Nucleic Acids Res* 37, 346-353.
- Tomari, Y., Matranga, C., Haley, B., Martinez, N., and Zamore, P.D. (2004). A protein sensor for siRNA asymmetry. *Science* 306, 1377-1380.
- Tseng, Y.C., Mozumdar, S., and Huang, L. (2009). Lipid-based systemic delivery of siRNA. *Adv Drug Deliv Rev* 61, 721-731.
- Ui-Tei, K., Naito, Y., Takahashi, F., Haraguchi, T., Ohki-Hamazaki, H., Juni, A., Ueda, R., and Saigo, K. (2004). Guidelines for the selection of highly effective siRNA sequences for mammalian and chick RNA interference. *Nucleic Acids Res* 32, 936-948.
- Ui-Tei, K., Naito, Y., Zenno, S., Nishi, K., Yamato, K., Takahashi, F., Juni, A., and Saigo, K. (2008). Functional dissection of siRNA sequence by systematic DNA substitution: modified siRNA with a DNA seed arm is a powerful tool for mammalian gene silencing with significantly reduced off-target effect. *Nucleic Acids Res* 36, 2136-2151.
- Urban-Klein, B., Werth, S., Abuharbeid, S., Czubyko, F., and Aigner, A. (2005). RNAi-mediated gene-targeting through systemic application of polyethylenimine (PEI)-complexed siRNA in vivo. *Gene Ther* 12, 461-466.
- van Rij, R.P., and Berezikov, E. (2009). Small RNAs and the control of transposons and viruses in *Drosophila*. *Trends Microbiol* 17, 163-171.

- Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A., *et al.* (2001). The sequence of the human genome. *Science* 291, 1304-1351.
- Vert, J.P., Foveau, N., Lajaunie, C., and Vandembrouck, Y. (2006). An accurate and interpretable model for siRNA efficacy prediction. *BMC Bioinform* 7, 520.
- Vickers, T.A., Koo, S., Bennett, C.F., Crooke, S.T., Dean, N.M., and Baker, B.F. (2003). Efficient reduction of target RNAs by small interfering RNA and RNase H-dependent antisense agents. A comparative analysis. *J Biol Chem* 278, 7108-7118.
- Vickers, T.A., Lima, W.F., Nichols, J.G., and Crooke, S.T. (2007). Reduced levels of Ago2 expression result in increased siRNA competition in mammalian cells. *Nucleic Acids Res* 35, 6598-6610.
- Vickers, T.A., Wyatt, J.R., and Freier, S.M. (2000). Effects of RNA secondary structure on cellular antisense activity. *Nucleic Acids Res* 28, 1340-1347.
- Wakefield, D.H., Klein, J.J., Wolff, J.A., and Rozema, D.B. (2005). Membrane activity and transfection ability of amphipathic polycations as a function of alkyl group size. *Bioconjug Chem* 16, 1204-1208.
- Walton, S.P., Stephanopoulos, G.N., Yarmush, M.L., and Roth, C.M. (1999). Prediction of antisense oligonucleotide binding affinity to a structured RNA target. *Biotechnol Bioeng* 65, 1-9.
- Walton, S.P., Stephanopoulos, G.N., Yarmush, M.L., and Roth, C.M. (2002). Thermodynamic and kinetic characterization of antisense oligodeoxynucleotide binding to a structured mRNA. *Biophys J* 82, 366-377.
- Wang, X., Xu, W., Mohapatra, S., Kong, X., Li, X., Lockey, R.F., and Mohapatra, S.S. (2008). Prevention of airway inflammation with topical cream containing imiquimod and small interfering RNA for natriuretic peptide receptor. *Genet Vaccines Ther* 6, 7.
- Watanabe, T., Totoki, Y., Toyoda, A., Kaneda, M., Kuramochi-Miyagawa, S., Obata, Y., Chiba, H., Kohara, Y., Kono, T., Nakano, T., *et al.* (2008). Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature* 453, 539-543.
- Weiss, J.N. (1997). The Hill equation revisited: uses and misuses, pp. 835-841.
- Weitzer, S., and Martinez, J. (2007). The human RNA kinase hC1p1 is active on 3' transfer RNA exons and short interfering RNAs. *Nature* 447, 222-226.
- Westerhout, E.M., and Berkhout, B. (2007). A systematic analysis of the effect of target RNA structure on RNA interference. *Nucleic Acids Res* 35, 4322-4330.

- Westerhout, E.M., Ooms, M., Vink, M., Das, A.T., and Berkhout, B. (2005). HIV-1 can escape from RNA interference by evolving an alternative structure in its RNA genome. *Nucleic Acids Res* 33, 796-804.
- Wolff, J.A., and Rozema, D.B. (2008). Breaking the Bonds: Non-viral Vectors Become Chemically Dynamic. *Mol Ther* 16, 8-15.
- Wurm, F.M. (2004). Production of recombinant protein therapeutics in cultivated mammalian cells. *Nat Biotechnol* 22, 1393-1398.
- Xiong, M.P., Bae, Y., Fukushima, S., Forrest, M.L., Nishiyama, N., Kataoka, K., and Kwon, G.S. (2007). pH-responsive Multi-PEGylated dual cationic nanoparticles enable charge modulations for safe gene delivery. *ChemMedChem* 2, 1321-1327.
- Yezhelyev, M.V., Qi, L., O'Regan, R.M., Nie, S., and Gao, X. (2008). Proton-sponge coated quantum dots for siRNA delivery and intracellular imaging. *J Am Chem Soc* 130, 9006-9012.
- Yoneyama, M., Kikuchi, M., Natsukawa, T., Shinobu, N., Imaizumi, T., Miyagishi, M., Taira, K., Akira, S., and Fujita, T. (2004). The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. *Nat Immunol* 5, 730-737.
- Yoshinari, K., Miyagishi, M., and Taira, K. (2004). Effects on RNAi of the tight structure, sequence and position of the targeted region. *Nucleic Acids Res* 32, 691-699.
- Zhang, S., Zhao, B., Jiang, H., Wang, B., and Ma, B. (2007). Cationic lipids and polymers mediated vectors for delivery of siRNA. *J Control Release* 123, 1-10.
- Zheng, X., and Bevilacqua, P.C. (2004). Activation of the protein kinase PKR by short double-stranded RNAs with single-stranded tails. *RNA* 10, 1934-1945.
- Zhu, X., and Galili, G. (2004). Lysine metabolism is concurrently regulated by synthesis and catabolism in both reproductive and vegetative tissues. *Plant Physiol* 135, 129-136.
- Zuhorn, I.S., Engberts, J.B., and Hoekstra, D. (2007). Gene delivery by cationic lipid vectors: overcoming cellular barriers. *Eur Biophys J* 36, 349-362.
- Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* 31, 3406-3415.

MICHIGAN STATE UNIVERSITY LIBRARIES



3 1293 03063 1307